Master Thesis

Analyzing the Indexing Effect on Stock Returns – Evidence from the German Stock Market

Faculty of Economics and Management
University of Lucerne

Supervisor
Prof. Dr. Lukas David Schmid

Author:	Jonas Furrer
E-Mail:	jonas.furrer@stud.unilu.ch
Address:	Eizmoosweg 5
	6330 Cham ZG
Student ID:	17-452-053
Submission Date:	17.05.2023

Abstract

We examine the indexing effect based on the DAX expansion from 30 to 40 companies, which happened in September 2021. Our chosen method is the newly developed generalized synthetic control method, as traditional methods fail or are limited by their assumptions in our context. We show the trajectory of the average treatment effect of the treated (ATT) 15 days before the announcement to 15 days after the introduction of the new provision (effective day). On the announcement day, we find an indexing effect of 0.95%. Furthermore, we notice a sharp increase in the ATT ahead of the announcement day, while the trading volume shows only a small spike in this period. Regarding the effective day, we identify a run-up by 2.7% to a peak in the ATT and a unique peak in trading volume, one day before the effective day. The effect has partially reversed within 15 days after the effective day. We find evidence of the selection criteria hypothesis, around the announcement day. Our findings on the announcement day agree with theories of the diminishing indexing effect. Around the effective day, we find evidence for the price pressure hypothesis. Together with its interpretation and explanations, our analysis expands the literature and deepens the understanding of indexing effects and the method to analyze them.

Table of content

List of figures

List of tables

List of abbreviations

AD	Announcement Day
ATT	Average Treatment Effect of the Treated
CAPM	Capital Asset Pricing Model
CAR	Cumulative Abnormal Return
DAX	Large-Cap German Stock Index
DiD	Difference in Difference
EBITDA	Earnings before Interest, Taxes, Depreciation and Amortization
ED	Effective Day
EPS	Earnings per Share
ETF	Exchange-Traded Fund
GSC	Generalized Synthetic Control
IAH	Investor Awareness Hypothesis
IH	Information Hypothesis
ISH	Imperfect Substitutes Hypothesis
LH	Liquidity Hypothesis
MDAX	Mid-Cap German Stock Index
OLS	Ordinary Least Squares
OTC	Over-the-Counter
PPH	Price Pressure Hypothesis
P/B	Price-to-Book Ratio
SCH	Selection Criteria Hypothesis
SDAX	Small-Cap German Stock Index
S&P	Standard and Poor’s
TecDAX	Technology German Stock Index

1 Introduction

The indexing effect describes an index premium (abnormal return) associated with the addition or deletion of a stock to or from a major index. The first papers on this indexing effect appeared in the late 1980s, focusing on composition changes in the S&P 500. To quantify the indexing effect, most researchers adopt the event study methodology, which is first used for this particular case by Shleifer (1986) and Harris and Gurel (1986). Some newer studies use more sophisticated approaches like the regression discontinuity design (Chang, Hong and Liskovich, 2015) or incorporate prediction models (Franz, 2020).

Numerous papers confirm the existence of the indexing effect and agree on the direction of the effect (positive abnormal return). However, economists are still debating possible explanations and have split themselves into two main argumentative divisions formulating hypotheses. On the one hand, we have demand-based hypotheses, which explain the effect through block purchases of index tracking funds and institutional investors trying to minimize tracking errors (Harris and Gurel, 1986; Lynch and Mendenhall, 1997). On the other hand, information-based hypotheses argue that indexing adds new information to the market, which is causing price changes (Dhillon and Johnson, 1991; Denis *et al.*, 2003). Besides the debate about the explaining factors, researchers find differing magnitudes, as there are numerous factors influencing stock prices. Given the existing literature, one can identify that the effect, measured in abnormal return, lies within a range of 1% to 7%. Key elements affecting the indexing effect are the transparency of the selection criteria and the selection process. Higher transparency implies more accurate index composition predictions by investors and thus, as a consequence, the effect is partially offset on the announcement day (AD) due to investors successively trading in advance trying to maximize returns (Petajisto, 2008; Fernandes and Mergulhão, 2016). With the increasing popularity of passive investing and the associated rise in index-tracking funds activity, also in Germany[1] (Deutsche Börse AG, 2022), the effect on the effective day (ED) should be magnified.

[1] Measured by the number of exchange-traded funds (ETF) and ETF trading volume on the German stock exchange XETRA.

A better understanding of the indexing effect is particularly interesting for market participants aiming to generate excess returns. From a scientific perspective, the indexing effect is connected to fundamental (behavioral-) economic theorems through its hypotheses. Understanding these hypotheses helps to verify existing theories about capital markets and the behavior of human participants.

We use the natural experiment in which the large-cap German stock index (DAX) was extended by 10 stocks. We analyze the effect of the cumulative abnormal return (CAR) using the Generalized Synthetic Control (GSC) method introduced by Xu (2017). With this method we address the problem that, in the pretreatment period, stocks to be included in the index often outperform stocks that will not be included. This is linked to performance-based inclusion criteria and follows the general idea that many explaining factors such as price-to-book (P/B) ratio, Profit-margins or firm size are better for firms that are more likely to be included. Due to the high level of transparency of the DAX inclusion criteria, we expect parts of the effect to happen before the AD. This is why our treatment starts 15 days ahead of the AD.

Given the new approach (GSC) applied to recent data of a completely new case, we can add to the scientific consensus regarding the direction of the effect, as well as shed some more light on the debate concerning the hypotheses and the magnitude of the effect. Furthermore, as the DAX has been rarely used for such analyses in the past, this paper can present new evidence about a topic extensively discussed in other geographic markets.

From here on, we will introduce the DAX index family and its regulations in subsection 1.1. Followed by an introduction to the existing literature and the affiliated hypotheses in section 2. In section 3, we show the basis of preparations and data used for the model in section 4. Closing in on section 5, we provide a detailed analysis of the results and elaborate on our findings as evidence of the mentioned hypotheses.

1.1 DAX index family and its regulations

The DAX is the most important equity index in Germany. It comprises the 40 biggest German companies (as of September 20, 2021), aiming to represent the German economy by containing approximately 80% of the market capitalization of the listed stocks (Deutsche Börse Group, no date a). Furthermore, it is also used as a benchmark for a large number of financial products. The mid-cap German stock index (MDAX) and the small-cap German stock index (SDAX) entail 50 mid-cap stocks and 70 small-cap stocks, respectively, while the technology German stock index (TecDAX) only focuses on companies in the technology sector (Deutsche Börse Group, no date a). STOXX Ltd. is the administrator of the DAX indices and is part of Qontigo, a subsidiary of the Deutsche Börse Group (Deutsche Börse Group, no date b).

The DAX entails the so-called “Blue Chip” stocks. Blue Chip companies are publicly traded and characterized as well-established, financially sound and stable with widely accepted products and services. To identify those companies, STOXX Ltd. publishes the “Guide to the DAX Equity Indices” multiple times a year. In this work, we use version 11.2.3 (STOXX Ltd., 2021) which lists the relevant criteria for the composition of the DAX and MDAX on our date of interest.

The accounting scandal of the DAX member Wirecard intensified a DAX composition reform (Storbeck, 2020). On October 5, 2020, STOXX Ltd. started a market consultation to revise the criteria of the DAX family to improve its quality and align it with international standards (Qontigo, 2020a). Experts considered the reform to be overdue (Gries, 2021). This is also shown by the high acceptance of the new rules by investors and other market participants (Qontigo, 2020b). The results were presented on November 24, 2020, and STOXX Ltd. published the new DAX-Family selection criteria, which would be introduced stepwise until September 2021 (Qontigo, 2020b).

The DAX Reform announcement stated that, in September 2021, the DAX would be extended by 10 companies to a total of 40 members to create a more representative Index (Qontigo, 2020b). The MDAX in return is reduced by ten companies to a total of 50 members. In other words, ten of the MDAX members are promoted to the DAX, which is a world premiere. On September 3, 2021, the names of the new members were publicized. Since March 2021, all companies of the DAX section indices have had to publish audited annual financial reports as well as quarterly statements (Qontigo, 2020b, p. 1). If the company does not meet the new standards, it will be removed from the DAX after a 30-day warning period. As a result, for a company, it is sufficient to be listed on the regular market and no longer to the Prime Standard of the Frankfurt Stock Exchange. This new rule allows for faster sanctions in case of violations. In addition, the company’s legal headquarters or operating headquarters have to be located in Germany, their shares must be continuously traded on the trading venue Xetra and the minimum free-float share is required to be 10% (STOXX Ltd., 2021, p. 26). Since March 2021, the new members also have been required to comply with the recommendation of the German Corporate Governance Code, forming an audit committee in the supervisory board (Qontigo, 2020b, p. 1). Companies without (with) index membership must show a minimum Order Book Volume of 1 bln (0.8) EUR over the last 12 months or a Turnover Rate[2] of 20% (10%), respectively (STOXX Ltd., 2021, p. 28). Since December 2020, all DAX candidates have been required to have positive earnings before interest, taxes, depreciation and amortization (EBITDA) for the two most recent fiscal years (Qontigo, 2020b, p. 1). Companies that fulfill all mentioned criteria will be ranked based on the Free Float Market Capitalization (STOXX Ltd., 2021, p. 29). The main reviews of potential index changes take place on a semiannual basis in March and September (Qontigo, 2020b, p. 1).

[2] The turnover rate is the 12 month order book volume divided by the market capitalization (as defined by STOXX Ltd. (2021))

2 Literature

2.1 Theoretical approaches

The following section introduces a variety of relevant literature and hypotheses. For a comprehensive overview, visit Appendix A.

2.1.1 Demand-based theories

The most prominent hypotheses summarized under demand-based theories are the imperfect substitutes hypothesis (ISH) and the price pressure hypothesis (PPH). These explain the indexing effect by applying the basic principle of supply and demand. The idea is that the price increase is driven by a forced increase in demand from investors, without any information-driven and intentional purchases. The EMH (Fama, 1970) suggests that all publicly available information is reflected in the stock price. Therefore, a stock inclusion provides no additional information to the market. It suggests that an investor cannot outperform the market by timing the market. Arguing based on the EMH, the effect on the ED is explained by upward pressure on the price of the stock (Harris and Gurel, 1986; Shleifer, 1986). The price pressure arises from index-tracking funds, since those funds must buy (sell) stocks added to (deleted from) the mentioned index near the ED to minimize the tracking error (Blume and Edelen, 2002). Further categorizing our hypotheses, economists also distinguish between permanent and temporary effects. According to the ISH, the indexing effect is permanent under the assumption that stocks have a long-term, downward-sloping demand curve (Shleifer, 1986). A price effect is only visible if no perfect substitutes exist. Otherwise, alternative financial products will be bought and the price remains unchanged. Such alternatives in theory are payoff-replications with other stocks, bonds or options. Although in practice, creating a perfect replication of an index is unlikely due to the nature of its market mass and various market frictions. Beneish and Whaley (1996), Lynch and Mendenhall (1997) and Wurgler and Zhuravskaya (2002) all find results supporting the ISH.

In contrast to the ISH, the PPH argues for temporary effects as it assumes short-run downward-sloping demand curves for stocks, while long-term demand is fully elastic. Following this hypothesis, stock prices will return to their original level once the short-term upward price pressure through index fund purchases eases (Harris and Gurel, 1986). Elliott and Warr's (2003) findings underpin the PPH (around the ED) by analyzing data from 1989 to 2000.

2.1.2 Information-based theories

When debating information-based theories, the information hypothesis (IH), investor recognition hypothesis (IRH) and investor awareness hypothesis (IAH), liquidity hypothesis (LH) and the selection criteria hypothesis (SCH) find most support from researchers. As the stock market is heavily reliant on the distribution of information, it is hard to derive the indexing effect from purely economic theories and control for potential confounders like news and investor recognition, since the media often makes index changes a topic of interest to the public (Merton, 1987; Barber and Odean, 2008; Hirshleifer, Lim and Theo, 2009). This leads to the information-based theories, implying that news about an addition (deletion) to (from) a relevant index puts a strong signal to the market. Dhillon and Johnson (1991), Cai (2007) and Jain (1987) find evidence of the IH implying that the inclusion in a major index is not an information-free event and could be a sign of their longevity and prospects of being an “industry leader” with elevated future performance. This is supported by findings from Denis et al. (2003), who find significant upward revisions in analysts’ earnings per share (EPS) forecasts for firms added to the S&P 500 index. The work of Gygax and Otchere (2010) shows as well that indexing is not an information-free event. Chen, Noronha and Singal (2004) counter the IH as they find no significant positive abnormal return before 1976 for additions to the S&P 500, suggesting that indexing is an information-free event. This hypothesis should cause a permanent indexing effect under the assumption that these investors will invest in such companies for the long term. Chan, Kot and Tang (2013) find in a long-term study that stocks deleted from an index outperform stocks added to an index.

Chen, Noronha and Singal (2004) find asymmetric price effects for S&P 500 additions (large effect) and deletions (small effect). They explain their findings with the IAH as their results are not consistent with the existing hypothesis. This theory builds on the investor recognition hypothesis (IRH) which suggests, from an investor’s perspective, the higher the awareness, the higher the chance of buying said stock (Merton, 1987). In contrast, it is difficult for investors to become unaware of a stock. Meaning that, after deletion, the stock is still prominent (Chen, Noronha and Singal, 2004). Consequently, the deletion effect is rather small in contrast to the addition effect and fades out over time as stocks slowly move from being present to being absent for investors[3] (Chen, Noronha and Singal, 2004). Furthermore, they suggest that the addition effect may be permanent, while the deletion effect is expected to be temporary if there is any effect at all (Chen, Noronha and Singal, 2004). The work of Elliott et al. (2006) and Chen, Shiu and Wei (2019) favor the IAH by analyzing the S&P 500 and an international study with 38 different indices.

The LH states that a positive addition effect is caused by increased liquidity (Chen, Noronha and Singal, 2004). This hypothesis originated in the work of Amihud and Mendelson (1986), who map returns as an increasing function of the bid-ask spread. As Chung et al. (1995) state, market makers set up the bid-ask spread according to the number of analysts. Since more analysts provide more detailed information about relevant stocks, the total amount of publicly available information increases, and therefore the information asymmetry between investors decreases. A decline in the information asymmetry leads to a reduction of the bid-ask spread, which in turn leads to higher trading volume and liquidity (Chung *et al.*, 1995). To complete the argumentation, the increased liquidity is followed by higher stock prices as the required rate of return for investors decreases (Shleifer, 1986). The paper by Hegde and McDermott (2003) supports this hypothesis by finding a significant association between higher CAR and the effective bid-ask spread during the event period. However, if it is assumed that index-tracking funds make large purchases and focus on the buy-and-hold-strategy, the share of the available float[4] will become smaller and liquidity will be reduced as a result. Becker-Blease and Paul (2006) provide another approach, explaining increased liquidity. They find that firms added to the S&P 500 register abnormal increases in capital expenditure, which is positively related to stock liquidity.

[3] The addition effect in the CAR was 3.171% and the deletion effect in the CAR was -1.168% at the announcement date with data in a time period between September 1976 and September 1989 (Chen, Noronha and Singal, 2004, pp. 1908-1909).
[4] $available\ float = outstanding\ shares - restricted\ shares - closely\ held\ shares$

The selection criteria hypothesis (Edmister, Graham and Pirie, 1994) argues that part of the effect is explained by the selection criteria for the index participation, as these criteria are often standard key performance indicators (KPIs) such as profitability, liquidity or size. The idea is that the same KPIs are also used by investors to fundamentally assess a company's value and make an investment decision. This leads to the natural consequence that better stocks are worth more and also more likely to be in an index. An improvement in these factors is then associated with the indexing effect (Bechmann, 2004). A key element influencing the SCH argumentation is the transparency of the selection criteria. Petajisto (2008) shows that due to the predictability of changes, there is a connection between transparency and magnitude. For this reason, Fernandes and Mergulhão (2016) analyze pretreatment periods and find effects before the announcement date (anticipatory effects) which explains up to 40% of the CAR. Franz (2020) shows that even though index changes are predictable, there is still an indexing effect where promoted firms generate up to a 1.42% 1-day premium and demoted firms generate 1.54% adverse returns.

2.2 Empirical work

To quantify the effect, most researchers have used the event-study approach. Shleifer (1986, p. 582) finds that stocks added to the S&P 500 index between 1976 and 1983 experienced a significant, positive risk-adjusted return, quantifying it at 2.79%. The effect occurs on the event day[5] and lasts for at least 10 days after the inclusion announcement. Harris and Gurel (1986, p. 821) investigate the indexing effect for the period of 1978 to 1983 and find an effect of 3.13%, which is nearly fully reversed after 2 weeks.

Since October 1989, Standard and Poor have changed their policy regarding index changes and have newly announced these 5 days in advance of the ED (Lynch and Mendenhall, 1997). With this change, the "S&P 500 Game" as described by Beneish and Whaley (1996) was born. Since index funds are forced to rebalance their portfolio near the effective date, risk-arbitrageurs speculate on price jumps due to block purchases

[5] Before 1989, the stock exchange announced changes to the S&P 500 on the evening after trading hours and these changes were implemented on the following day.

and thus buy ahead of the rebalancing. Beneish and Whaley (1996) find an index addition effect of ~3% on the AD and up to 7% on the ED, with a permanent effect of 5%. Lynch and Mendenhall (1997) and Wurgler and Zhuravskaya (2002) identify similar patterns to Beneish and Whaley (1996).

The longevity of the indexing effect and the effect itself continue to heat up debates among researchers, causing new researchers from different countries to join the pool. This sheds new light on the historic S&P 500 studies, as different markets (indices) may show disparate results. Mazouz and Saadouni (2007), Mase (2007) and Vespro (2006) study the FTSE 100 and find evidence of the PPH, while Fernandes and Mergulhão (2016) favor the ISH. Analyzing China's CSI 300 index, Wang, Murgulov and Haman (2015) find results supporting the PPH. This contradicts most studies focusing on the S&P 500, which argue with the ISH (Shleifer, 1986; Dhillon and Johnson, 1991; Beneish and Whaley, 1996). Analyzing the MSCI Standard Country Indices for 29 countries, Chakrabarti et al. (2005) find similar index effects for the US stock market with additional mild evidence of the PPH and LH in the UK and Japan.

Contrary to the evaluated average indexing effect of 1% to 7% on the AD and ED, Wang, Murgulov and Haman (2015, p. 22) find a significant abnormal return effect on the AD of 0.537% followed by a significant effect of -0.312% on the ED. Only 55.3% of the analyzed stocks show a positive abnormal return, in contrast to Chen, Noronha and Singal's (2004) analysis of the S&P 500, where 94% exhibit positive abnormal returns (Wang, Murgulov and Haman, 2015, pp. 21 and 23). For index deletions, they find negative abnormal returns of -0.398%. Mase (2007, p. 467) identifies a cumulative abnormal return of -0.1% on the AD by analyzing composition changes in the FTSE 100 from June 1992 to March 2005. Looking at the promotions from a small cap to a large cap FTSE index, Biktimirov and Li (2014, no pagination) find a permanent increase in stock prices with a one-time effect of 0.8% in CAR on the AD.

2.3 Further implications

As noted in section 1.1, the DAX-composition is strictly rule-based. The rule book is publicly available and clearly specifies the date and the reasoning for composition changes, which is why we define the DAX to be highly transparent. Although the basic criteria are public, the S&P 500 is considered to be less transparent, as the final decision is made by a committee behind closed doors and the revision dates are unknown (S&P Dow Jones Indices, 2023). Furthermore, the committee can set exceptions for the said criteria. This leads to an aggravation of the composition predictability. We recognize that different indices have different rules for index membership and that rule transparency is particularly important, which has been explained under the SCH. As we have established that the DAX is highly transparent, this stretches the one-day theoretical magnitude of the effect over a period before the AD (and ED). In addition, note that the effect is bigger for companies which have no index membership than firms who move between indices (Shankar and Miller, 2006).

The majority of researchers use data from periods around the year 2000. Given that capital markets have grown tremendously in total capitalization and market dynamics as well as participants have changed, the question arises whether the indexing effect itself changes over time. This possible time inconsistency is addressed by some researchers as part of the index tracking system in the context of the PPH and ISH. As Basak and Pavlova (2013) point out, the price pressure, caused by institutions, boosts index stock prices. This means that, as more index-tracking funds arise, the indexing effect on the ED should be bigger. On the contrary, a rise in hedge funds is associated with more efficient markets and more shorting (Hanson and Sunderam, 2014) which can reduce the effect. Soe and Dash (2008, p. 6) compare the effect for the time period from September 1998 to August 2003 and September 2003 to August 2008, finding that the effect for the S&P 500 shrinks from 6.05% to 3.76%. They also identify a declining pattern for the DAX 30. Beneish and Whaley (1996) expect the S&P game to disappear due to index funds that rebalance during the announcement interval. A more recent study by Chang, Hong and Liskovich (2015) confirms the idea of a diminishing indexing effect by analyzing the Russel 2000 from 1996 to 2012 using a regression discontinuity design, which is supported by Mecca (2022), who finds that the decline of the effect continues by examining data from 1993-2021.

3 Data and model preparation

All data and R-Code is available on my GitHub repository: https://github.com/JonasFur/Master-Thesis.

3.1 Data sources

For our analysis, we need data on daily stock prices, index prices and trading volume, which we obtain from Yahoo! Finance (https://finance.yahoo.com) and TradingView (https://de.tradingview.com). The two sources are used to verify the correctness of the data. Since we are working with closing prices and ignore dividends, we focus on the DAX price index and not the performance index. The daily trading volume describes the number of shares that are bought or sold on a trading day. We take the data from TradingView. Starting from July 1, 2020, to three weeks after the ED, October 8, 2021, we end up with 326 observations for each company.

In our descriptive statistics, we compare the contestants of the treatment and control group. Our period of interest for the fundamental data is the first three quarters of 2021. The fundamentals are market capitalization, P/B ratio, EBITDA margin, dividend yield and basic EPS, of which the first two are expressed weekly and the latter quarterly. The main source is TradingView (https://de.tradingview.com). The data were manually completed and verified with the official financial reports of the individual companies as well as the Macrotrends website (https://www.macrotrends.net). We used data from Yahoo! Finance (https://finance.yahoo.com) for the sector affiliation, as well as the Frankfurt stock exchange (https://www.boerse-frankfurt.de). For more detailed information about the data sources see Bibliography extension A/B.

3.2 Data preparation

In a first step, we calculate the daily returns using equation (1):

$$r_{it} = \frac{p_{it} - p_{it-1}}{p_{it-1}} \tag{1}$$

In which r_{it} is the return on a share i at time t and p_{it} (p_{it-1}) denotes the price of a share on day t (the trading day before time t).

As we want to analyze the abnormal returns generated following an index inclusion, we decide to calculate the expected return with the market model. The market model is widely accepted in the literature. It incorporates the individual firm risk according to the capital asset pricing model (CAPM) as presented in the work of Lintner (1965) and Sharpe (1964). With the following ordinary least squares (OLS) regression design, we estimate the parameters $\hat{\alpha}$ and $\hat{\beta}$ for each stock:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it} \tag{2}$$

Where α_i denotes the firm-specific intercept and β_i is the coefficient for each firm, a measure of the sensitivity of the firm's returns on the market return. r_{mt} denotes the market return at time t and ε_{it} marks the error term for each firm at time t. With the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$, the fitted return $\hat{r}_{it}$ for each stock at time t is calculated. As the following formula denotes, the individual abnormal return for time t (AR_{it}) is the difference between the observed excess return and the fitted return:

$$AR_{it} = r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{mt} \tag{3}$$

The cumulative abnormal return for each stock at time t (CAR_{it}) is calculated using the following formula:

$$CAR_{it} = \sum AR_{it} \tag{4}$$

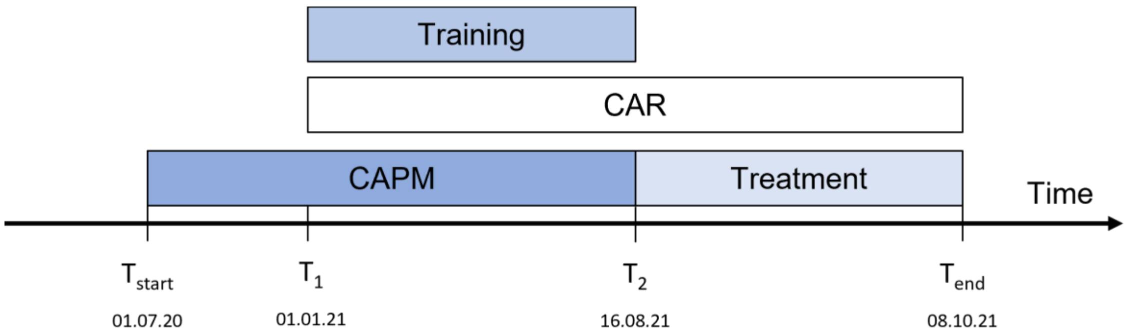
3.3 Time frame of the model

Figure 1 provides a short graphical summary of the relevant periods in our setting. The CAPM period starts at $T_{start,}$ namely on July 1, 2020, and ends at T_2, on August 16, 2021, and is used to estimate the stock-specific parameters with the OLS model. We choose a period that contains enough data to estimate reliable parameters and exclude gross market anomalies such as the Covid-19 crash in early 2020 (see Appendix B). In addition, we make sure that our CAPM period does not overlap with the treatment period, otherwise there would be event effects from the treatment period included in our CAPM coefficients. As Xu (2017) mentions, when the number of control units and the training period increases, the bias of the average treatment effect of the treated (ATT) of the

GSC approaches 0[6]. Our training period starts on January 1, 2021, and ends at T_2. In this period, we train our model and calculate the weights to predict the counterfactuals. We need as much data as possible for this period to incorporate post-crisis market fluctuations without having to exclude many companies from the analysis due to previous index composition changes. We calculate the CAR starting at T_1 until 15 days after the effective date, October 8, 2021, (T_{end}). The treatment period is our period of interest, starting 15 days before the AD, T_2, and ending on October 8, 2021, 15 days after the ED (T_{end}). This period is common in the literature and allows us to identify preannouncement effects and the duration of the effect (temporary or permanent).

The official statement concerning the companies eligible for the DAX expansion was published on Friday, September 3, 2021, after the closing bell. Consequently, on Monday, September 6, 2021, is the first trading day with information about the new index composition. As there is no accessible data on Over-the-Counter (OTC) trading, we do not incorporate it in our analysis. For this reason, we define September 6, 2021, as the AD and denote this as day 0. We begin our analysis 15 days before the AD. September 20, 2021, is the first day with the new index composition and is defined as the ED (day 10). October 8, 2021, is the last day of our analysis marked as day 24.

Figure 1: Model timeline



Note: This figure shows the different periods on a timeline, which are important for the basic structure of the model. The CAPM period is used to estimate the model parameters. The CAR period is used to calculate the CAR. The training period is used to train our model and finally apply the model in the treatment period.
Source: Own figure.

[6] For a more detailed description, see Xu (2017).

3.4 Treatment and control group

For information about the MDAX and DAX composition and their changes, we use the DAX and MDAX Composition Reports[7] quarterly appraisal from STOXX Ltd. Our treatment group consists of the ten companies which switch from the MDAX to the DAX. As we train our model starting at T_1, Porsche Automobil Holding will be excluded because it enters the MDAX after T_1. Our initial control group includes the 50 members of the MDAX according to the MDAX Composition Report of September 3, 2021[7], after the deduction of our treatment group members. We exclude five firms from our analysis which will drop to the SDAX at the ED to exclude counter-effects caused by demotions. Furthermore, we exclude Beiersdorf AG and Auto1 Group SE as they enter the MDAX after T_1. Due to the lack of data, we have to exclude Lufthansa AG from the control group. After the review in December 2020, none of the relevant companies were newly included in the MDAX. We can therefore assume that we have no interfering indexing effects from the past. In short, our treatment group consists of 9 units and our control group consists of 42 units. For a complete list of our treatment and control group, see Appendix C.

3.5 Descriptive Statistics

In this section, we take a closer look at the companies of interest. Table 1 provides an overview of the data for the control group and the treatment group as well as the two groups combined. Looking at the total market capitalization (3.) there is a wide range from 1.49 bln € to 94.36 bln €, with 75% of firms being below 11.1 bln €. We detect a significant difference between the control (1.) and treatment group (2.), as the mean of the treatment group is more than 20 bln € higher than the mean of the control group. Furthermore, the treatment group reports a higher range than the control group. For the P/B ratio, we take the natural logarithm as we simply want to analyze whether the groups are undervalued $\left(\log\left(\frac{P}{B}\right) < 0\right)$ or overvalued $\left(\log\left(\frac{P}{B}\right) > 0\right)$ and if there is a numerical difference between the groups. Both groups are overvalued but we note a significant difference showing that firms in the treatment group are more overvalued

[7] https://www.dax-indices.com/de/web/dax-indices/zusammensetzung

(4./5.). Interestingly, all members of the treatment group are overvalued without exceptions and only the control group contains undervalued companies. Furthermore, we are looking at the profitability measured by the EBITDA margin, calculated by dividing EBITDA by revenue. Profitability ranges from highly unprofitable to highly profitable firms with an average of 20.79% (9.), with most observations achieving an EBITDA margin between 10.5% and 30.23%. Although the mean values are close to each other (7./8.), the profitability of the control group is extremely stretched, which suggests sectoral or operating differences and makes comparability harder. Row (12.) shows the total mean dividend yield of 1.53%, a measure to compare dividend payments. We obtain a relatively wide range while the distribution is skewed to the right, especially for the treatment group (11.). Comparing the treatment and control group, the treatment group pays significantly less dividends relative to the share price. Although we observe a wide range in the EPS, the mean and median of the two groups lie close to each other, with the variation mainly caused by control units. For the first period CAR [-15,0], the treatment group outperforms the control group by far (16./17.) with 1.41% vs 0.08%. In the following period between the AD and ED [0,10], the treatment group reports a CAR of -1.20%, eliminating most of the previous period's overperformance compared to the control group, which remained unchanged (19./20.). In the final period after the ED [10,24], both groups show significant negative CAR at -1.03% for the control group and -4.08% for the treatment group (22./23.).

Table 1: Descriptive statistic

	Mean	Median (Min, Max)	Q1, Q3	SD	N
Market Cap. (bln €)					
1. Control Group	6.29	5.53 (1.49, 19.51)	3.50, 7.38	3.85	1638
2. Treatment Group	28.74	15.77 (8.34, 94.36)	12.72, 35.00	23.99	351
3. Total	10.25	6.34 (1.49, 94.36)	3.80, 11.10	13.67	1989
Log P/B Ratio (%)					
4. Control Group	0.98	0.90 (-1.46, 3.71)	0.28, 1.75	1.01	1638
5. Treatment Group	2.08	2.02 (1.02, 3.50)	1.41, 2.47	0.69	351
6. Total	1.18	1.12 (-1.46, 3.71)	0.39, 2.02	1.04	1989
EBITDA Margin (%)					
7. Control Group	21.15	18.08 (-127.42, 69.29)	11.27, 29.76	21.89	120
8. Treatment Group	18.75	11.64 (2.23, 39.80)	8.58, 35.04	13.13	21
9. Total	20.79	18.05 (-127.42, 69.29)	10.50, 30.23	20.80	141
Dividend Yield (%)					
10. Control Group	1.76	1.53 (0.00, 7.53)	0.47, 2.90	1.61	111
11. Treatment Group	0.46	0.06 (0.00, 1.85)	0.00, 0.52	0.72	24
12. Total	1.53	1.45 (0.00, 7.53)	0.04, 2.43	1.57	135
Basic EPS (€)					
13. Control Group	0.76	0.45 (-13.01, 13.00)	0.22, 0.90	2.02	126
14. Treatment Group	0.69	0.49 (-0.03, 2.38)	0.41, 0.89	0.54	24
15. Total	0.75	0.47 (-13.01, 13.00)	0.24, 0.90	1.86	150
CAR (%) [-15,0]					
16. Control Group	0.08	-0.05 (-11.59, 11.24)	-2.19, 2.67	3.95	630
17. Treatment Group	1.41	0.38 (-6.26, 13.14)	-1.22, 3.54	4.30	135
18. Total	0.32	0.00 (-11.59, 13.14)	-1.99, 2.70	4.04	765
CAR (%) [0,10]					
19. Control Group	0.10	0.19 (-10.59, 7.89)	-1.74, 2.17	3.09	420
20. Treatment Group	-1.20	-0.63 (-10.46, 3.99)	-2.49, 0.59	2.76	90
21. Total	-0.13	0.01 (-10.59, 7.89)	-1.79, 1.91	3.07	510
CAR (%) [10,24]					
22. Control Group	-1.03	-0.74 (-49.71, 16.21)	-3.50, 2.04	6.65	630
23. Treatment Group	-4.08	-2.94 (-21.56, 4.79)	-5.95, 0.52	5.93	135
24. Total	-1.57	-1.32 (-49.71, 16.21)	-3.93, 1.54	6.63	765

Note: This table presents the characteristics of the analyzed companies for the control and treatment group as well as the two groups in total, measured in the first 3 quarters of 2021 (row 1. – 15.). It should be noted that some companies in the 7. – 15. rows are excluded from the analysis due to missing data. Market cap and P/B ratio were collected weekly while EBITDA margin, dividend yield and basic EPS were collected quarterly. Rows 16. - 24. show the daily CAR, measured over different periods, indicated in the squared brackets. The columns show key attributes of the data as well as the number of observations (N).
Source: Own table. Data: https://de.tradingview.com, Bibliography extension A/B

It should be noted that there are sectorial and structural differences that can have a different impact on the KPIs shown above. Table 2 shows that 9 different industries are represented across the analyzed companies. From the control group, at least one company is represented in each industry, while only 4 of the 9 industries are represented in the treatment group. As a consequence, the share per industry between the groups is different, which can be seen for e.g. consumer cyclical or healthcare. The Basic Materials sector is also overrepresented in the treatment group (3.), but the Industrials sector is more strongly represented in the control group than in the treatment group (4.).

Table 2: Sector analysis

Sector	**Control Group (N=42)**	**Treatment Group (N=9)**	**Total (N=51)**
1. Consumer Cyclical	3 (7.1%)	3 (33.3%)	6 (11.8%)
2. Healthcare	4 (9.5%)	3 (33.3%)	7 (13.7%)
3. Basic Materials	6 (14.3%)	2 (22.2%)	8 (15.7%)
4. Industrials	9 (21.4%)	1 (11.1%)	10 (19.6%)
5. Communication Services	6 (14.3%)	0 (00.0%)	6 (11.8%)
6. Financial Services	3 (7.1%)	0 (00.0%)	3 (05.9%)
7. Real Estate	5 (11.9%)	0 (00.0%)	5 (09.8%)
8. Technology	5 (11.9%)	0 (00.0%)	5 (09.8%)
9. Utilities	1 (2.4%)	0 (00.0%)	1 (02.0%)

Note: This table shows the composition of the control and treatment group by sector, expressed in absolute and percentage terms for the control and treatment group as well as the total. In addition, the number of group members is shown in the header.
Source: Own table. Data: https://finance.yahoo.com, https://www.boerse-frankfurt.de

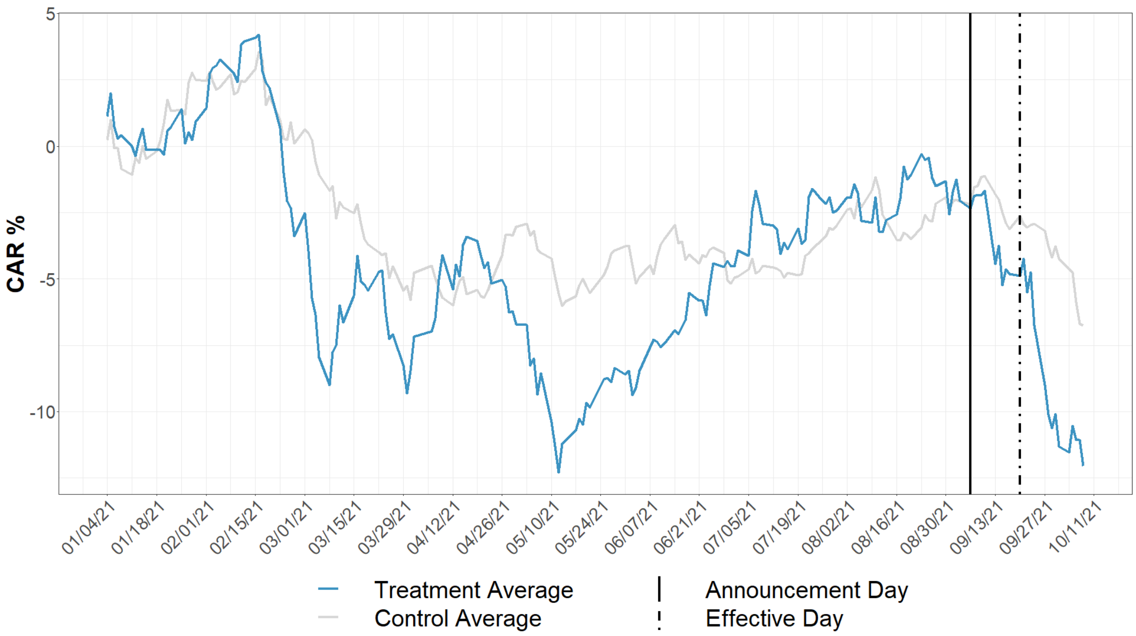
Overall, the KPIs show acceptable differences which should not interfere with our method, despite a significant difference in market capitalization, as this was expected due to the nature of the inclusion criteria. Concluding on the descriptive analysis, we should not face destructive issues in the application and interpretation of the GSC, considering differences between the treatment and control group. Interestingly, the realized CAR for the treatment group is inferior to the control group, which opposes our and the literature's expectations.

4 Method

4.1 Short overview

The literature on the indexing effect almost exclusively uses the event study. Our one-time setting, where 10 companies are admitted from the MDAX to the DAX at the same time, allows for a more timely method to determine the indexing effect. The difference in difference (DiD) method is often used to make causal inferences from panel data. As Angrist and Pischke (2009) say, there are only non-biased causal estimates in parallel worlds. The DiD method is based on strong identification assumptions, as in the absence of the intervention the treatment group follows a similar trend as the control group in the post-treatment period. Due to the data, the parallel trend assumption does not hold for the pretreatment period and therefore we infer that parallel trends are implausible for the post-treatment period (see Figure 2).

Figure 2: Average CAR of the treatment and control group



Note: This figure shows the average daily CAR in percent for the treatment group (blue) and the control group (grey). The CAR was calculated according to section 3.2. The vertical black line marks the AD (September 6, 2021), while the vertical dot-dashed line marks the ED (September 20, 2021).
Source: Own figure. Data: https://finance.yahoo.com

Xu (2017) proposes a method using time-series cross-sectional-data to estimate the ATT when the parallel trend assumption is unlikely to hold. He also addresses the problem of unobserved, time-varying confounders. He combines the method of matching within the synthetic control method, developed by Abadie (2005), Abadie, Diamond and Hainmueller (2010, 2015) and Abadie and Gardeazabal (2003), and the IFE model proposed by Bai (2009) and unifies the synthetic control method with linear fixed effects. In addition, we take advantage of the fact that this method is applicable to cases with multiple treated units as our setting proposes.

Therefore, we apply the GSC method following the paper of Xu (2017). We are using control group information from the training period to impute counterfactuals for each treated unit for the treatment period. The ATT is calculated as the difference between the observed average outcome of the treated and the mean of the imputed counterfactuals of the treated. For our analysis, we use the R-Package "gsynth", developed and published by Xu and Liu (2022). The next section takes a closer look at Xu's (2017) model.

4.2 Methodology

4.2.1 Basics

Following Xu (2017), we define the number of units in the treatment group as N_{tr} and the total number of units in the control group as N_{co}. Hence, the total number of units is N, while τ and ς are defined as the sets of units in the treatment and the control group. We have 157 pretreatment periods defined as T_0, with a starting point of treatment for the treated at (T_0 +1), observed for 40 days q, as in our setting all units of the treatment group are treated at the same time. The control group is never exposed to treatment during the entire period T of 197 days. As Xu (2017) suggests, let us assume that our variable of interest (CAR) is given by a linear factor model:

$$CAR_{it} = \delta_{it}D_{it} + x'_{it}\beta + \lambda'_i f_t + \varepsilon_{it} \quad (5)$$

D_{it} is equal to 1 if the unit is part of the treatment group and in a period after T_0, otherwise D_{it} equals 0. δ_{it} denotes the heterogeneous treatment effect on unit i at time t. $x'_{it}\beta$ is defined as a vector of observed covariates and their unknown parameters. As we do not include any covariates in our analysis, we remove this term from equation (5). ε_{it} denotes the error term of unit i at time t and $\lambda_i' f_t$ is the factor component of the model and takes a linear additive form as in equation (6).

$$\lambda_i' f_t = \lambda_{i1} f_{1t} + \lambda_{i2} f_{2t} + \cdots + \lambda_{ir} f_{rt} \tag{6}$$

It covers the conventional additive unit and time-fixed effects as special cases. The subscript r denotes the number of unobserved common factors. According to Xu (2017), this expression can include all unobservable random variables, as long as they can be decomposed into a multiplicative form (i.e. $U_{it} = a_i b_t$).

Defining $CAR(1)$ and $CAR(0)$ as the potential outcomes for $D = 1$ and $D = 0$ for unit i at time t, $\delta_{it} = CAR_{it}(1) - CAR_{it}(0)$ for any $i \in \tau$ and $t > \mathrm{T}_0$. This leads to equation (7):

$$CAR_i = D_i \circ \delta_i + F\lambda_i + \varepsilon_i, i \in 1,2, \dots N_{co}, N_{co} + 1, \dots, N \tag{7}$$

In which $CAR_i = [CAR_{i1}, CAR_{i2}, \dots, CAR_{iT}]'$, $D_i = [D_{i1}, D_{i2}, \dots, D_{iT}]'$, $\delta_i = [\delta_{i1}, \delta_{i2}, \dots, \delta_{iT}]'$ and $\epsilon_i = [\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{iT}]'$ are vectors with the dimension of $(T \times 1)$ and $F = [f_1, f_2, \dots, f_T]'$ is given as a $(T \times r)$ matrix, while r denotes the number of unobserved factors. The stacked outcome of all units in the control group is given in equation (8).

$$CAR_{co} = F\Lambda_{\mathrm{co}}' + \varepsilon_{co} \tag{8}$$

Since $\Lambda_{co} = \left[\lambda_1, \lambda_2, \dots, \lambda_{N_{co}}\right]'$ is a $(N_{co} \times r)$ matrix the product $F\Lambda_{co}'$ is a $(T \times N_{co})$ matrix as well as $\varepsilon_{co} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{N_{co}}]$ and $CAR_{co} = \left[CAR_1, CAR_2, \dots, CAR_{N_{co}}\right]$. According to Bai (2003, 2009), Xu (2017) normalizes all factors and sets the constraint that they are orthogonal to each other ($F'F/T = I_r$ and $\Lambda_{\mathrm{co}}'\Lambda_{\mathrm{co}} = diagonal$).

Our variable of interest is the average treatment effect of the treated, which shows the average effect of all units included in the DAX after T_0. The ATT for each time period is calculated as the difference in the average of the CAR for each unit i at time t treated and the average of the CAR for each unit i at time t not treated as stated in equation (9).

$$ATT_{t,t>T_0} = \frac{1}{N_{tr}} \sum_{i \in \tau} [CAR_{it}(1) - CAR_{it}(0)] = \frac{1}{N_{tr}} \sum_{i \in \tau} \delta_{it} \tag{9}$$

Like Abadie, Diamond and Hainmueller (2010), Xu (2017) treats the treatment effects δ_{it} as given, conditional on the data sample. This means that we do not predict $CAR_{it}(0)\ i \in\ \tau$, but treat it as if it were missing data. To calculate an estimator for the treatment effect of each treated unit we need to estimate $\hat{F}, \widehat{\Lambda}_{co}, \hat{\lambda}_i$ and the estimated counterfactual ($\widehat{CAR}_{it}(0)$). Using an out-of-sample prediction method based on the work of Bai (2009), Xu (2017) provides three steps to estimate the $\widehat{CAR}_{it}(0)$. In the first step, an IFE model is estimated using data of the control group. Hence, the estimates $\hat{F}\ and\ \widehat{\Lambda}_{co}$ are obtained through equation (10).

$$(\hat{F}, \widehat{\Lambda}_{co}) = \underset{\tilde{F}, \widetilde{\Lambda}_{co}}{argmin} \sum_{i \in \varsigma} (CAR_i - \tilde{F}\tilde{\lambda}_i)'\ (CAR_i - \tilde{F}\tilde{\lambda}_i) \tag{10}$$

These minimizations are conducted under the constraints $\tilde{F}'\tilde{F}/T = I_r$ and $\widetilde{\Lambda}_{\mathrm{co}}'\widetilde{\Lambda}_{co} = diagonal$. The factor loadings for each treated unit are calculated by using the estimated variables from equation (10) by minimizing the prediction mean squared error in the pretreatment periods (11). The superscript “0” indicates that only data from the pretreatment period is used.

$$\begin{aligned} \hat{\lambda}_i &= \underset{\tilde{\lambda}_i}{argmin} (CAR_i^0 - \hat{F}^0\tilde{\lambda}_i)'(\ CAR_i^0 - \hat{F}^0\tilde{\lambda}_i) \\ &= \left(\hat{F}^{0\prime}\hat{F}^0\right)^{-1}\hat{F}^{0\prime}CAR_i^0, i \in \tau \end{aligned} \tag{11}$$

Finally, we are able to estimate the counterfactual of the treated using the estimated variables $\hat{F}$ and $\hat{\lambda}_i$:

$$\widehat{CAR}_{it}(0) = \hat{\lambda}_i'\hat{f}_t, i\ \in\ \tau, t > T_0 \tag{12}$$

We estimate the ATT by using equation (12) with the assumption that $CAR_{it}(0) = \widehat{CAR}_{it}(0)$.

4.2.2 Selecting the Model

In the next section, we show the five-step procedure by Xu (2017), how to choose the number of factors r using the leave-one-out-cross-validation method. Starting with a given r, an IFE model is estimated based on data from the control group and estimates $\hat{F}$. In step two, a cross-validation-loop iterates over all pretreatment periods, leaving out data of one data point $s\ \in \{1, \dots, T_0\}$ for the treatment units. Using all other data of the pretreatment period, an OLS regression estimates the factor loadings ($\hat{\lambda}_{i,-s}$) for each treated unit. The subscript $-s$ denotes all data from the pretreatment period except for the time period s. Next, the outcome of the missing time period is estimated for each unit and is compared to the observed outcome at s, calculating the prediction error. In the third step, the mean squared prediction error (MSPE) is calculated with the given r. Now we repeat the first three steps with different values for r and choose the r* that minimizes the MSPE. Xu (2017) conducts a Monte Carlo analysis, showing that this method performs well (even in small data sets – not our concern). In our work we set the r = 22, which does not minimize the MSPE, however, we set a limit to include even more factors because of the computational power. In addition, the MSPE could only be minimally improved from r = 22 onward.

4.2.3 Inferences

Lastly, we want to obtain the uncertainty estimates and therefore a confidence interval for our $\widehat{ATT}_t$. To do so, we estimate the conditional variance of the ATT estimator (i.e. $Var_{\varepsilon}(\widehat{ATT}_t|D, \Lambda, F)$). As ε_i is the only random variable that is not conditioned on, we can interpret it as the unexplained variation of the outcome unrelated to treatment assignment. Xu (2017) provides a method including a parametric bootstrap procedure via resampling residuals based on the resampling scheme:

$$\begin{aligned} \widetilde{CAR}_i(0) &= \hat{F}\hat{\lambda}_i + \tilde{\varepsilon}_i, \forall i \in \varsigma \\ \widetilde{CAR}_i(0) &= \hat{F}\hat{\lambda}_i + \tilde{\varepsilon}_i^p, \forall i \in \tau. \end{aligned} \tag{13}$$

Where $\widetilde{CAR}_i(0)$ denotes the simulated outcomes without the treatment. The estimated conditional mean is based on $\hat{F}\hat{\lambda}_i$. $\tilde{\varepsilon}_i$ and $\tilde{\varepsilon}_i^p$ are the resampled residuals for each unit of the control group or the treatment group, respectively. As $\hat{F}$ is estimated using only data of the companies not getting included in the DAX, this estimation fits better for a unit from the control group resulting usually in a higher variance of $\tilde{\varepsilon}_i^p$ than $\tilde{\varepsilon}_i$. Since these are from different statistical distributions, we can interpret $\tilde{\varepsilon}_i$ as the in-sample error of the IFE model and $\tilde{\varepsilon}_i^p$ as the prediction error of the model.

In the first step, a unit from the control group is selected and is treated as if it received the treatment in time $t > T_0$. The rest of the control group gets resampled with replacements to get a full control group with the size N_{co}. Now a GSC is applied to this sample to obtain the residual (i.e. $\hat{\varepsilon}_{(m)}^{p} = CAR_i - \widehat{CAR}_i(0)$) and store them in the vector $\hat{e}^{p}$. To obtain enough observations for the confidence interval, we run a loop B_1 times (i.e. $m \in \{1, \dots, B_1\}$). In step two, a GSC method is applied to the original data to estimate the $\widehat{ATT}_t$ for all $t > T_0$, the coefficient, the fitted values, and the residuals of the control units and store the values in the vector $\widehat{CAR}_{co}$ and $\hat{e}$. In step 3, a bootstrap loop runs B_2 times ($B_2 = 2000$), using a sample $S^{(k)}$ while the vector of $\tilde{\varepsilon_\iota}$ and $\tilde{\varepsilon}_j^{p}$ are randomly selected from e and e^{p}. $\widehat{CAR}_i(0) = \hat{F}\hat{\lambda}_i$ and hence we sample $S^{(k)}$ by:

$$\begin{aligned} \widehat{CAR}_i^{(k)}(0) &= \widehat{CAR}_i(0) + \tilde{\varepsilon}_i, i \in \varsigma \\ \widehat{CAR}_i^{(k)}(0) &= \widehat{CAR}_i(0) + \tilde{\varepsilon}_i^{p}, j \in \tau \end{aligned} \tag{14}$$

At this point, the treatment effects are not part of the simulated treated counterfactuals. Now, the GSC method is applied to the samples $S^{(k)}$. By adding the $\widehat{ATT}_t$ from step two to the newly obtained ATT from the $S^{(k)}$ samples we obtain the bootstrap estimate $\widehat{ATT}_t^{(k)}$ for all $t > T_0$. According to all these obtained average treatment effects, the variance is calculated and the confidence intervals are constructed.

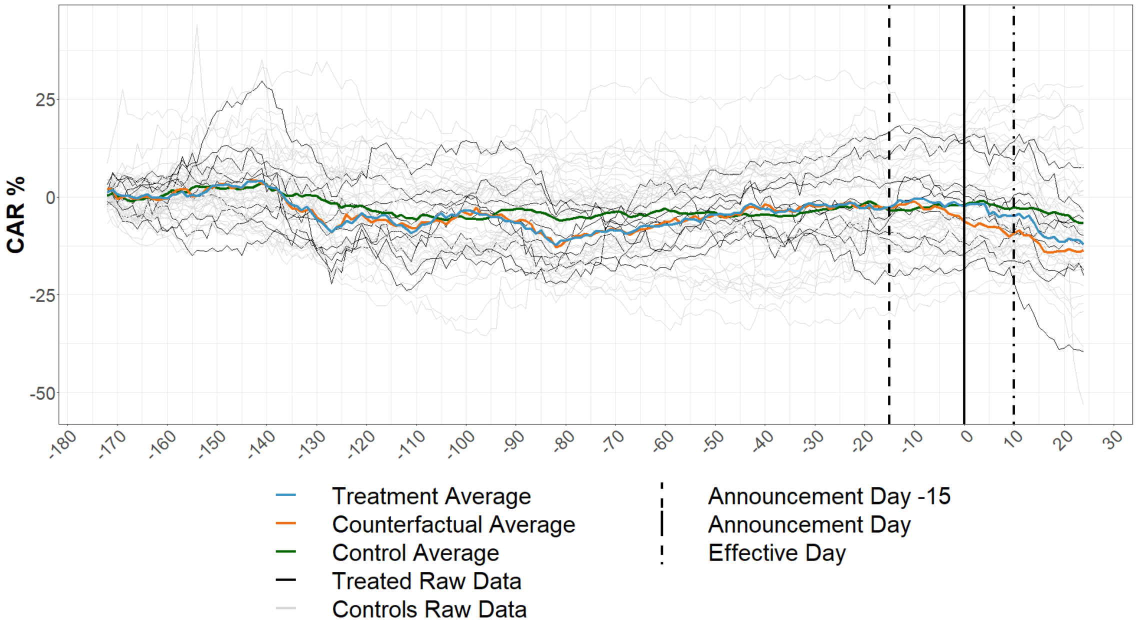
4.3 Limitations

Although the GSC method has many advantages for our setting, there are also limitations. Our GSC is based on our outcome variable and the unobserved factors captured by this approach, without including any control variables. The question arises as to how well the GSC could define the unobserved factors and whether control variables remove bias. Pickett, Hill and Cowan (2022) show in their paper that control variables could improve the RMSE, but both models with and without control variables are very close to each other relative to other methods. Therefore, we assume that our result is very similar to the result of a model that includes control variables.

As Xu (2017, p. 59) states, the number of pretreatment periods must exceed the number of unit fixed effects, otherwise the "incidental parameters" can lead to bias in the estimated treatment effects. In our case, this should not be an issue, as our analysis includes 157 pretreatment periods and only 42 unit fixed effects. While the GSC method can capture unobserved variables that can be decomposed into a multiplicative form as described in section 4.2.1, it cannot include unobserved confounders that are independent across units. The following issue may lead to erroneous conclusions. First, we need to consider, unlike the synthetic control method, that the GSC model will impute the counterfactuals despite large differences in the treatment and control group regarding common factors. The synthetic factors in the factor model are representative of covariates such as firm fundamentals which, in reality, explain stock returns. The GSC initially ignores this practical relationship and forces common factors (parallel trends in CAR), which implies that the two groups must report factors in common territory. Our visual diagnostic test confirms that the CAR of the treatment group lies within the range of the control group's CAR (Figure 3). Following economic theory on stock dynamics, this implies common explanatory factors. In addition, our descriptive analysis further supports the existence of common factors. It is therefore possible to label the generated synthetic treatment unit as correct.

Another limitation of our setting is the external validity. Many of the studies that examine the indexing effect use larger control and treatment groups. In addition, the effect is measured at different points in time, by incorporating index-composition-changes over several years. This makes the result less susceptible to time-specific events. It is well known that the stock market is susceptible to a lot of different factors, such as policy changes or sector-specific news. In our setting, we do not control for such events, which may lead to a bias in our estimator if time-specific events in $t > T_0$ have different influences on our treatment and control group. Furthermore, the question arises, whether our results are comparable to the results in the literature. As in the literature, we analyze the stock addition to a major index. Contrary to most literature, our treatment group units were members of another well-established index, namely the MDAX. Furthermore, our setting includes a DAX expansion and is not a normal DAX composition change. However, included companies may receive more attention because a DAX expansion has never taken place. This may affect our estimator, depending on the explanation described in section 2.1. As the new rules and their implementation dates are clearly communicated in 2020, the predictability is not impaired relative to a normal composition change.

Figure 3: Visual diagnostic test



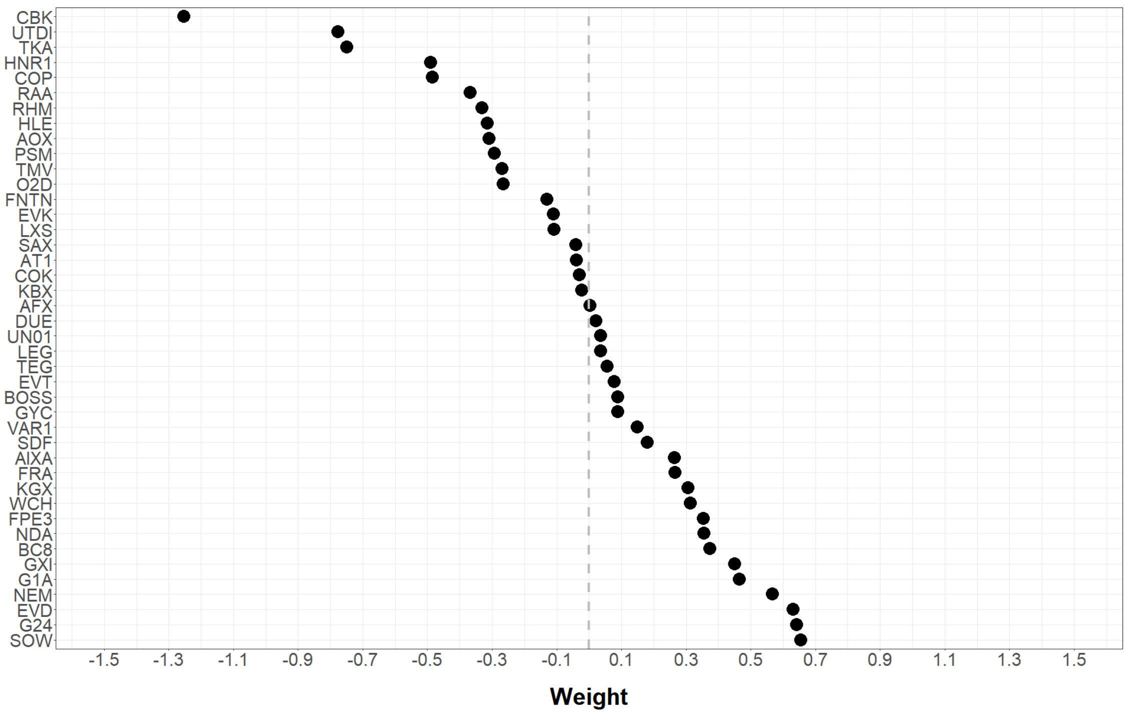
Note: This figure shows each company's daily CAR in percentage, divided into control group (light grey) and treatment group (black). The CAR was calculated according to section 3.2. The blue line shows the unweighted percentage mean of the treatment group, while the orange line shows the mean of the counterfactuals calculated by GSC. The black dashed vertical line marks day AD -15 (August 16, 2021) and stands for the transition from the training period to the treatment period. The black vertical line shows the AD (September 6, 2021) and the black dot-dashed line shows the ED (September 20, 2021). The description of the vertical lines is valid for all following figures. There are 197 periods in total, starting at -172 to 24.
Source: Own figure. Data: https://finance.yahoo.com

5 Results

A note on the structure of section 5: To align the theory and our findings, we present the findings stepwise and match them to the hypotheses in an attempt to avoid confusion. Results from supplementary analyses (volume) are directly integrated into the text.

Before we look at the imputed counterfactuals, we want to have a look at the unit weights. Figure 4 shows the weights for the different companies ordered from small to large. About 85% of the weights are in the range of -0.5 and 0.5, while the largest weight is still below 0.7. The most negative weighting is given to Commerzbank with a rounded -1.25. The counterfactual is therefore not calculated using only some units, as is often the case with the standard synthetic control method when the units show large differences[8].

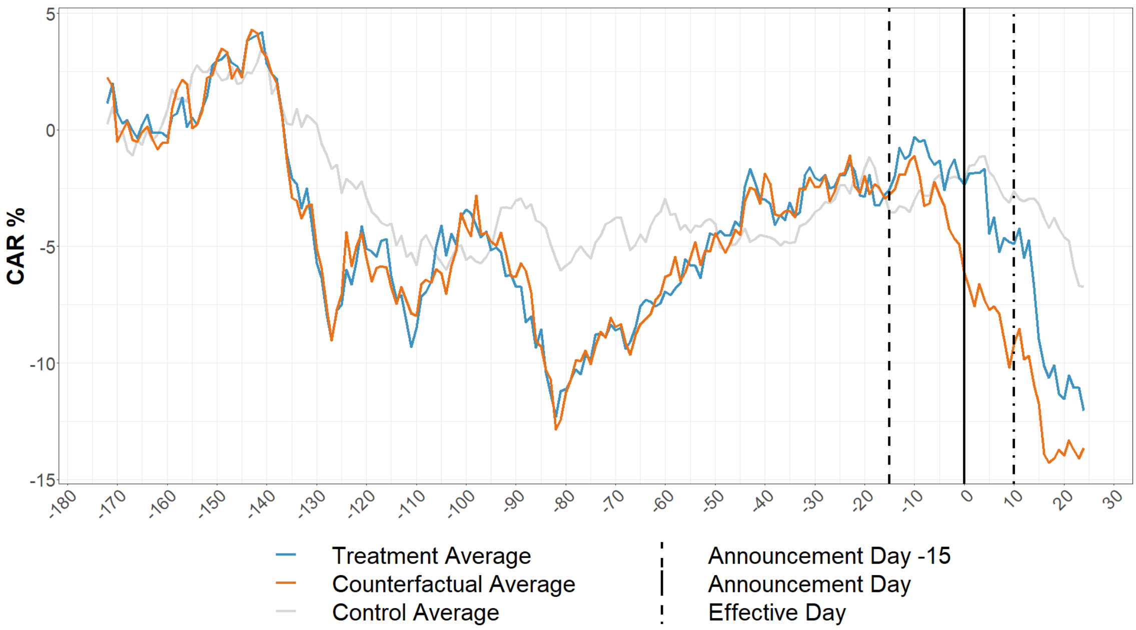
Figure 4: Unit weights



Note: This figure shows the weighting (black points) of the companies from the control group, according to the calculations of the GSC model, sorted by weighting. The average weight is almost 0.
Source: Own figure. Data: https://finance.yahoo.com

[8] With the synthetic control method, a weighting of 0 is possible, which means that the control group is reduced.

Figure 5: Average CAR of the treated and its counterfactual



Note: This figure shows the average CAR in percent for the treatment group (blue) as well as its counterfactual (orange). These are calculated on the basis of the returns, see section 3.2. The grey line marks the average CAR in percent for the control group and serves as a comparison.
Source: Own figure. Data: https://finance.yahoo.com

Next, we present our preliminary results by plotting the CAR of the different groups over time (see Figure 5). The blue line describes the average observed CAR of the treatment group, while the orange line represents the calculated counterfactual. The grey line represents the equally weighted average CAR of the control group. The counterfactual represents the synthetic treatment group, which does not receive treatment ($CAR(0)$). With a training-period MSPE of 3.45%, the GSC has managed to create a well-fitting synthetic treatment group. Up until 4 days ahead of the AD, the trajectories of the counterfactual and treatment unit do not diverge. Whereas from AD-4 onwards, the counterfactual decreases to a much lower level compared to the treatment group stagnating at ~-2%. It is also interesting to note that on the AD, the treatment group shows a slight increase, while the counterfactual continues to decline sharply.

After the ED, the two groups show a similar development. After a peak on ED+1, both groups fall sharply. On day 16, the counterfactual finds its bottom, whereas the treatment group continues to slightly decrease. These results indicate that without the treatment, the treatment group would have fallen more, especially around the AD.

However, it can also be seen that after the AD and ED event, the treatment group decreases more relative to the counterfactual, which hints at an effect reversal.

The difference between the average CAR of the treated and the average CAR of the counterfactuals denotes our ATT. Sub-figure 6a shows the progression of the calculated ATT in percent over time, while the 0 on the x-axis represents the AD. The grey area defines the 95% confidence interval as calculated in section 4.2.3. We can see that this interval occupies a large spectrum in the treatment period, which means that the true ATT lies within this spectrum with a certainty of 95%. The large confidence interval is most likely caused by our small sample size, which suggests being cautious when discussing related outputs. The red dotted line denotes the largest and the smallest value of the calculated ATT (-2.2%, 2.1%) in the training period. We assume that everything in this range is part of the natural variation of the ATT and thus cannot be identified as an effect attributable to the treatment. In addition, we examine the trading volume of our treatment group as only those are relevant for our interpretation. We calculate a treatment group volume average of 0.398 million using data from the training period. In sub-figure 6b, we plot the trading volume on the ATT for the treatment period to compare them. For an overview of the analyzed period, see Appendix D.

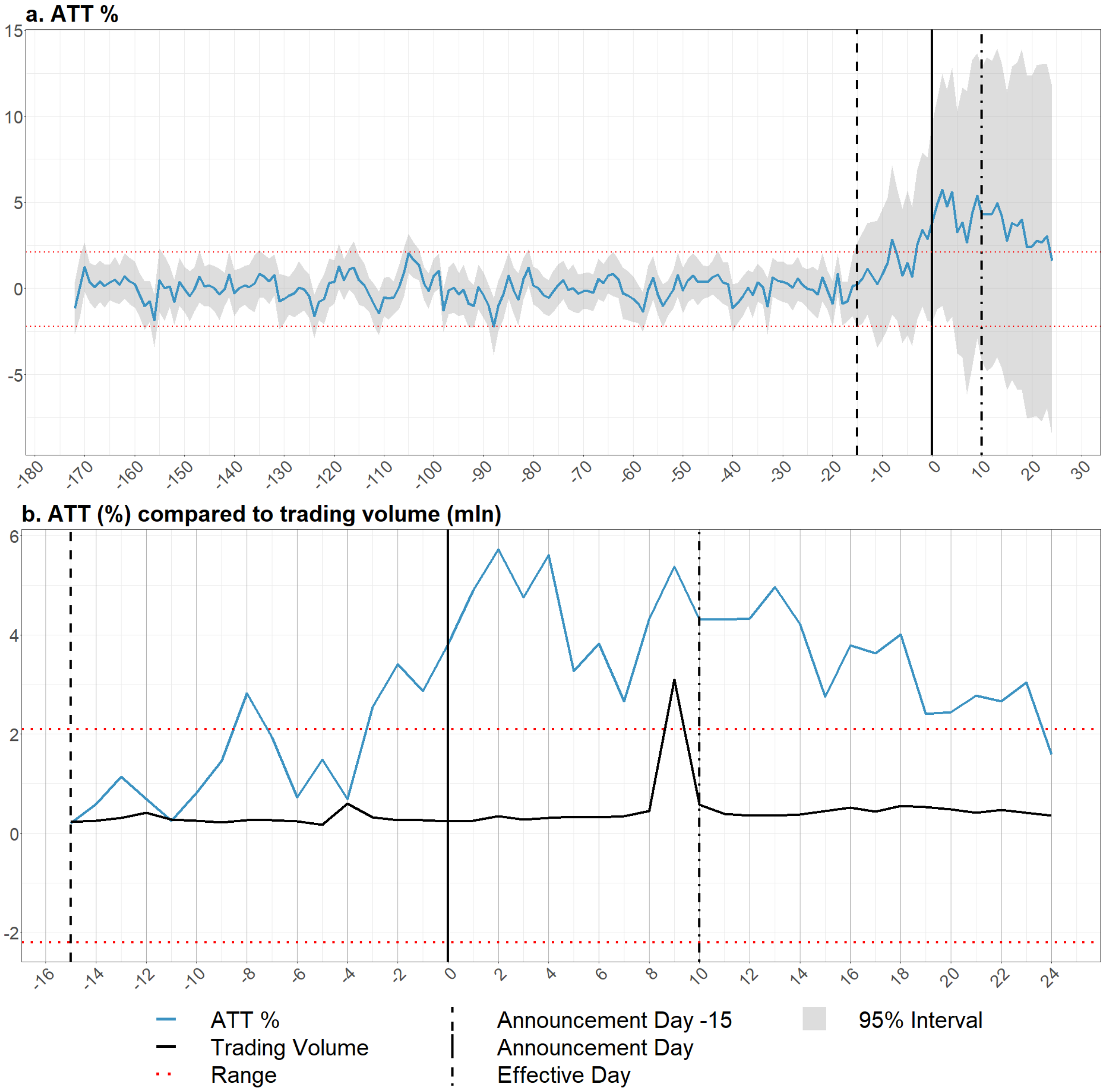
In sub-figure 6a, for the period 15 days before the AD, we find two spikes in the ATT that exceed the threshold of 2.1%. The first peak on day -8 is an unexpected event and has no literary context. Hence, this could be some random event (e.g. news) or a model-specific issue, which is unlikely to be the case. Furthermore, the ATT drops sharply after day -8, below the 2.1% threshold. For these reasons, we neglect this peak. Up to four days before the AD, we see only a small divergence between the CAR of the counterfactual and the observed treatment group and therefore conclude that there is no effect of the treatment on the treatment group (ATT) in this period. The second peak, however, two days before the AD (starting with the rise on day -4), exceeds the 3% mark. Additionally, 4 days in advance of the AD, we note a small spike in trading volume of 0.566 million. This is the last day relevant for the re-evaluation of the firms and the composition change. At this time, arbitrageurs can now with certainty predict the composition change and buy relevant stocks at the end of that day, which explains this peak. From the EMH's point of view, it is less intuitive why investors buy these shares before the AD, as there should be no effect. However, we do observe an effect around the AD, which is an argument countering the EMH. Given that the rules and selection procedures are highly transparent and we find preannouncement effects and the volume peak, we attribute this to the SCH. In our case, we thus conclude that it was possible to predict which companies would be included in the DAX. Incorporating this information into the trading strategy, more shares of these companies were bought by alleged arbitrageurs before the announcement. Our findings also hint at the IAH, although to correctly attribute the effect to the IAH, we also need to further analyze the exclusion effect. According to the PPH, our identified big increase in the ATT before the AD should be met with an equivalent increase in the trading volume. As can be seen in sub-figure 6b, this is not the case. Therefore, we conclude that the PPH is not a suitable explanation for the indexing effect on the AD.

The increase in the ATT from day -1 to day 0 is approximately 0.95%. However, the ATT increases by almost 3% until two days after the AD and reaches its peak of a total of 5.7% CAR. Our results tend to be at the lower end of the observed effects of the literature. This is consistent with the idea that transparency reduces the magnitude due to the stated arguments regarding arbitrageurs. Furthermore, this reduced magnitude is consistent with the theory that index membership before the treatment reduces the indexing effect. Given that our effect is smaller than the historical average, our results are also consistent with the theory on the diminishing indexing effect (section 2.3).

However, after AD+4, we see a reversal of the effect and the ATT drops to 2.7%, three days before the ED, followed by a sharp increase with a next peak shortly before the ED. At this point, the trading volume shows extremely abnormal behavior and increases to 3.1 million (see Sub-figure 6b). According to the literature on demand-based hypotheses, we attribute this peak to portfolio rebalancing by institutional investors (index-tracking funds), as those investors try to minimize the tracking error through big purchases of the relevant stocks, normally one day before the index composition change. These block purchases create price pressure that could not be absorbed by the trading volume and therefore explain parts of the effect. However, as shown, the ATT rises even before these purchases. We attribute this result to the "S&P game" described by Beneish and Whaley (1996), in which arbitrageurs buy the relevant stocks before the block purchases of the index-tracking funds and then sell them to the institutional investors at a premium.

When interpreting the results of the indexing effect according to their existing hypotheses, the persistence of the observed effect is important. Between the AD and the ED, we note a sharp decrease in the ATT. This trend is interrupted by an effect caused by the ED as described above. Hence, we are unable to make a clear statement on whether a part of the effect is permanent. The effect decreases after the ED but does not reverse completely. The effect reversal after the ED indicates the existence of temporary price effects, which is likely caused by the block purchases of index funds in connection with the change in the composition of the DAX. These findings point towards the PPH, which states that the composition change has a temporary effect on stock prices, triggered by institutional investors. Considering the ISH, IH and LH we cannot fully exclude these hypotheses with confidence, as these would require a full reversal of the effect. Whereas, in our analysis, we have only identified a downward-sloping trend after the ED. However, we find that 15 days after the ED, there is still a much smaller but visible effect, which can be seen in sub-figure 6b by the ATT closing below the upper bound. In conclusion, this means that the ISH, IH and LH need further investigation. In addition, an exclusion of the IAH is not possible because we do not control for variables that could increase investors' awareness, nor analyze index deletions and thus cannot determine the asymmetry of the effects of additions and deletions.

Figure 6: ATT and trading volume



Note: Figure a. shows the ATT in percentage (blue) over the entire training and treatment period, while the grey area shows the 95% confidence interval calculated as described in section 4.2.3. The red dotted lines mark the maximum and minimum values of the ATT of the training period, located at 2.1 and -2.2. The black dashed vertical line marks day AD -15. The black vertical line shows the AD and the black dot-dashed line shows the ED.
Figure b. compares the average number of shares sold or bought on a trading day in million for the treatment group (black) and the calculated ATT (blue) in percent over the treatment period. The trading volume is the equally weighted average of the volume, while the ATT was calculated using the GSC method. The y-axis applies to the ATT and the trading volume.
Source: Own figure. Data: https://finance.yahoo.com

6 Conclusion

In this paper, we investigate the indexing effect around the AD and the ED using the DAX expansion back in September 2021. We find a preannouncement effect, starting to rise 4 days ahead of the AD, with a small spike in trading volume. These findings hint at the SCH, although our interpretation is limited as parts of the effect may still be caused by the IAH. We find an ATT increase on the AD of 0.95%, which is on the lower end of the effect described in the existing literature and consistent with the hypothesis of the diminishing indexing effect. Starting 4 days ahead of the ED, the ATT rises about 2.7% and peaks one day in front of the ED. The trading volume shows a big peak the day ahead of the ED, as index-tracking funds normally rebalance their portfolio one day ahead of the composition change. These findings hint at the “S&P game”, as arbitrageurs buy the affected stocks before the rebalancing of the index fund providers, which triggers an effect before the ED and thus can be seen in an increase in the ATT. After the ED, the effect shows a downward-sloping trend, meaning a reversal of the effect, which is attributed to the PPH. 15 days after the ED, there is still a visible effect, even though it shrinks below the maximum ATT of the training period. Therefore, we cannot exclude the ISH, IH and LH for certain.

The comparability of our results is restricted by limited external validity. Furthermore, the results are limited by the fluctuations of the ATT in the training period as small effects cannot be interpreted. Moreover, we have to interpret the results with caution, since the calculated confidence interval has a wide range. Further research could consider more detailed trading volume analyses using the GSC and additionally examine index changes from the DAX to the MDAX to go into more detail on the various hypotheses, especially the IAH. For even more precise results, covariates like firm fundamentals and momentum could be included in the GSC model. In addition, research with even larger sample sizes over different points in time could be conducted to improve external validity and comparability to previous research. Our paper extends the literature on the indexing effect by applying a new method to recent data from the DAX index.

7 Bibliography

References

Abadie, A. (2005) 'Semiparametric Difference-in-Differences Estimators', *The Review of Economic Studies*, 72(1), pp. 1–19. doi: 10.1111/0034-6527.00321.

Abadie, A., Diamond, A. and Hainmueller, J. (2010) 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program', *Journal of the American Statistical Association*, 105(490), pp. 493–505. doi: 10.1198/jasa.2009.ap08746.

Abadie, A., Diamond, A. and Hainmueller, J. (2015) 'Comparative Politics and the Synthetic Control Method', *American Journal of Political Science*, 59(2), pp. 495–510. doi: 10.1111/ajps.12116.

Abadie, A. and Gardeazabal, J. (2003) 'The Economic Costs of Conflict: A Case Study of the Basque Country', *American Economic Review*, 93(1), pp. 113–132. doi: 10.1257/000282803321455188.

Amihud, Y. and Mendelson, H. (1986) 'Asset pricing and the bid-ask spread', *Journal of Financial Economics*, 17(2), pp. 223–249. doi: 10.1016/0304-405X(86)90065-6.

Angrist, J. D. and Pischke, J.-S. (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bai, J. (2003) 'Inferential Theory for Factor Models of Large Dimensions', *Econometrica*, 71(1), pp. 135–171. doi: 10.1111/1468-0262.00392.

Bai, J. (2009) 'Panel Data Models With Interactive Fixed Effects', *Econometrica*, 77(4), pp. 1229–1279. doi: 10.3982/ECTA6135.

Barber, B. M. and Odean, T. (2008) 'All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors', *Review of Financial Studies*, 21(2), pp. 785–818. doi: 10.1093/rfs/hhm079.

Basak, S. and Pavlova, A. (2013) 'Asset Prices and Institutional Investors', *American Economic Review*, 103(5), pp. 1728–1758. doi: 10.1257/aer.103.5.1728.

Bechmann, K. L. (2004) 'Price and Volume Effects Associated with Changes in the Danish Blue-Chip Index: The KFX Index', *Multinational Finance Journal*, 8(1/2), pp. 3–34. Available at: https://ssrn.com/abstract=2625140 (Accessed: 16 May 2023).

Becker-Blease, J. R. and Paul, D. L. (2006) 'Stock Liquidity and Investment Opportunities: Evidence from Index Additions', *Financial Management*, 35(3), pp. 35–51. doi: 10.1111/j.1755-053X.2006.tb00146.x.

Beneish, M. D. and Whaley, R. E. (1996) 'An Anatomy of the "S&P Game": The Effects of Changing the Rules', *The Journal of Finance*, 51(5), pp. 1909–1930. doi: 10.1111/j.1540-6261.1996.tb05231.x.

Biktimirov, E. N. and Li, B. (2014) 'Asymmetric stock price and liquidity responses to changes in the FTSE SmallCap index’, *Review of Quantitative Finance and Accounting*, 42(1), pp. 95–122. doi: 10.1007/s11156-012-0335-7.

Blume, M. E. and Edelen, R. M. (2002) *On Replicating the S&P 500 Index*, *Rodney L. White Center for Financial Research Working Paper No. 08-02*. Pennsyvania. Available at: https://ssrn.com/abstract=315545.

Cai, J. (2007) ‘What’s in the News? Information Content of S&P 500 Additions’, *Financial Management*, 36(3), pp. 113–124. doi: 10.1111/j.1755-053X.2007.tb00083.x.

Chakrabarti, R. *et al.* (2005) ‘Price and volume effects of changes in MSCI indices – nature and causes’, *Journal of Banking & Finance*, 29(5), pp. 1237–1264. doi: 10.1016/j.jbankfin.2004.04.002.

Chan, K., Kot, H. W. and Tang, G. Y. N. (2013) ‘A comprehensive long-term analysis of S&P 500 index additions and deletions’, *Journal of Banking & Finance*, 37(12), pp. 4920–4930. doi: 10.1016/j.jbankfin.2013.08.027.

Chang, Y.-C., Hong, H. and Liskovich, I. (2015) ‘Regression Discontinuity and the Price Effects of Stock Market Indexing’, *Review of Financial Studies*, 28(1), pp. 212–246. doi: 10.1093/rfs/hhu041.

Chen, H.-L., Shiu, C.-Y. and Wei, H.-S. (2019) ‘Price effect and investor awareness: Evidence from MSCI Standard Index reconstitutions’, *Journal of Empirical Finance*, 50, pp. 93–112. doi: 10.1016/j.jempfin.2019.01.002.

Chen, H., Noronha, G. and Singal, V. (2004) ‘The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation’, *The Journal of Finance*, 59(4), pp. 1901–1930. doi: 10.1111/j.1540-6261.2004.00683.x.

Chung, K. H. *et al.* (1995) ‘Production of information, information asymmetry, and the bid-ask spread: Empirical evidence from analysts’ forecasts’, *Journal of Banking & Finance*, 19(6), pp. 1025–1046. doi: 10.1016/0378-4266(94)00068-E.

Denis, D. K. *et al.* (2003) ‘S&P 500 Index Additions and Earnings Expectations’, *The Journal of Finance*, 58(5), pp. 1821–1840. doi: 10.1111/1540-6261.00589.

Deutsche Börse AG (2022) *ETF TRADING ON XETRA - Facts & figures for the year 2021*. Frankfurt am Main. Available at: https://www.xetra.com/resource/blob/2925940/97038c491fed1eae903271530ace62d7/data/Factsheet-Zahlen-und-Fakten-2021_en.pdf (Accessed: 27 March 2023).

Deutsche Börse Group (no date a) *DAX – benchmark and barometer for the German economy*. Available at: https://deutsche-boerse.com/dbg-en/media/deutsche-boerse-spotlights/spotlight/DAX-benchmark-and-barometer-for-the-German-economy-139948 (Accessed: 6 April 2023).

Deutsche Börse Group (no date b) *STOXX indices*. Available at: https://www.deutsche-boerse.com/dbg-en/products-services/ps-indices-analytics/ps-stoxx-indices (Accessed: 6 April 2023).

Dhillon, U. and Johnson, H. (1991) ‘Changes in the Standard and Poor’s 500 List’, *The*

Journal of Business, 64(1), pp. 75–85. Available at: https://www.jstor.org/stable/2353073.

Edmister, R. O., Graham, A. S. and Pirie, W. L. (1994) 'Excess Returns of Index Replacement Stocks: Evidence of Liquidity and Substitutability', *Journal of Financial Research*, 17(3), pp. 333–346. doi: 10.1111/j.1475-6803.1994.tb00196.x.

Elliott, W. B. *et al.* (2006) 'What Drives the S&P 500 Inclusion Effect? An Analytical Survey', *Financial Management*, 35(4), pp. 31–48. doi: 10.1111/j.1755-053X.2006.tb00158.x.

Elliott, W. B. and Warr, R. S. (2003) 'Price Pressure on the NYSE and Nasdaq: Evidence from S&P 500 Index Changes', *Financial Management*, 32(3), pp. 85–99. Available at: https://www.jstor.org/stable/3666385.

Fama, E. F. (1970) 'Efficient Capital Markets: A Review of Theory and Empirical Work', *The Journal of Finance*, 25(2), pp. 383–417. doi: 10.2307/2325486.

Fernandes, M. and Mergulhão, J. (2016) 'Anticipatory effects in the FTSE 100 index revisions', *Journal of Empirical Finance*, 37, pp. 79–90. doi: 10.1016/j.jempfin.2016.02.009.

Franz, F.-C. (2020) 'Forecasting index changes in the German DAX family', *Journal of Asset Management*, 21(2), pp. 135–153. doi: 10.1057/s41260-020-00153-6.

Gries, L. (2021) *Warum die DAX-Familie größer wird*, *tagesschau*. Available at: https://www.tagesschau.de/wirtschaft/finanzen/dax-40-erweiterung-leitindex-boerse-101.html (Accessed: 17 January 2023).

Gygax, A. F. and Otchere, I. (2010) 'Index composition changes and the cost of incumbency', *Journal of Banking & Finance*, 34(10), pp. 2500–2509. doi: 10.1016/j.jbankfin.2010.04.007.

Hanson, S. G. and Sunderam, A. (2014) 'The Growth and Limits of Arbitrage: Evidence from Short Interest', *Review of Financial Studies*, 27(4), pp. 1238–1286. doi: 10.1093/rfs/hht066.

Harris, L. and Gurel, E. (1986) 'Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures', *The Journal of Finance*, 41(4), pp. 815–829. doi: 10.1111/j.1540-6261.1986.tb04550.x.

Hegde, S. P. and McDermott, J. B. (2003) 'The liquidity effects of revisions to the S&P 500 index: an empirical analysis', *Journal of Financial Markets*, 6(3), pp. 413–459. doi: 10.1016/S1386-4181(02)00046-0.

Hirshleifer, D., Lim, S. S. and Theo, S. H. (2009) 'Driven to Distraction: Extraneous Events and Underreaction to Earnings News', *The Journal of Finance*, 64(5), pp. 2289–2325. doi: 10.1111/j.1540-6261.2009.01501.x.

Jain, P. C. (1987) 'The Effect on Stock Price of Inclusion in or Exclusion from the S&P 500', *Financial Analysts Journal*, 43(1), pp. 58–65. doi: 10.2469/faj.v43.n1.58.

Lintner, J. (1965) 'Security Prices, Risk, and Maximal Gains From Diversification', *The*

Journal of Finance, 20(4), pp. 587–615. doi: 10.2307/2977249.

Lynch, A. W. and Mendenhall, R. R. (1997) 'New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index', *The Journal of Business*, 70(3), pp. 351–383. doi: 10.1086/209722.

Mase, B. (2007) 'The Impact of Changes in the FTSE 100 Index', *The Financial Review*, 42(3), pp. 461–484. doi: 10.1111/j.1540-6288.2007.00179.x.

Mazouz, K. and Saadouni, B. (2007) 'The price effects of FTSE 100 index revision: what drives the long-term abnormal return reversal?', *Applied Financial Economics*, 17(6), pp. 501–510. doi: 10.1080/09603100600690085.

Mecca, G. J. (2022) *The Influence of Exchange-Traded Funds in the Index Inclusion Effect*. Doctoral Thesis, Rollins College, Crummer Graduate School of Business. Available at: https://scholarship.rollins.edu/dba_dissertations/40 (Accessed: 13 May 2023).

Merton, R. C. (1987) 'A Simple Model of Capital Market Equilibrium with Incomplete Information', *Journal of Finance*, 42(3), pp. 483–510. doi: 10.1111/j.1540-6261.1987.tb04565.x.

Petajisto, A. (2008) 'Selection of an Optimal Index Rule for an Index Fund', *SSRN Electronic Journal*. doi: 10.2139/ssrn.1264698.

Pickett, R. E. M., Hill, J. and Cowan, S. K. (2022) *The Myths of Synthetic Control: Recommendations for Practice, Cash Transfer Lab NYU Working Paper. New York*. Available at: https://as.nyu.edu/content/dam/nyu-as/cashtransferlab/documents/RPJHSC_SyntheticControl_041122.pdf.

Qontigo (2020a) *Qualification Criteria, International Standards, Transparency: Market Consultation on DAX Reform Launched*. Available at: https://qontigo.com/market-consultation-on-dax-reform-launched/ (Accessed: 13 May 2023).

Qontigo (2020b) *Results of Market Consultation*. Zug. Available at: https://www.stoxx.com/document/Resources/MarketConsultation/Results_of_Market_Consultation_DAX_Index_ENG.pdf (Accessed: 6 April 2023).

S&P Dow Jones Indices (2023) *S&P U.S. Indices Methodology*. Available at: https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-us-indices.pdf (Accessed: 3 April 2023).

Shankar, S. G. and Miller, J. M. (2006) 'Market Reaction to Changes in the S&P SmallCap 600 Index', *The Financial Review*, 41(3), pp. 339–360. doi: 10.1111/j.1540-6288.2006.00146.x.

Sharpe, W. F. (1964) 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance*, 19(3), pp. 425–442. doi: 10.1111/j.1540-6261.1964.tb02865.x.

Shleifer, A. (1986) 'Do Demand Curves for Stocks Slope Down?', *The Journal of Finance*, 41(3), pp. 579–590. doi: 10.1111/j.1540-6261.1986.tb04518.x.

Soe, A. M. and Dash, S. (2008) ‘The Shrinking Index Effect: A Global Perspective’, *SSRN Electronic Journal*. doi: 10.2139/ssrn.1568122.

Storbeck, O. (2020) ‘Deutsche Börse pushes for Dax overhaul after Wirecard debacle’, *Financial Times*, 5 October. Available at: https://www.ft.com/content/713c4986-aa45-4248-9c4d-e46355443fba (Accessed: 17 January 2023).

STOXX Ltd. (2021) *Guide to the DAX Equity Indices*. Available at: https://www.dax-indices.com/document/News/2021/August/Guide%20to%20the%20DAX%20Equity%20Indices_V11.2.3_valid30.08.2021.pdf (Accessed: 16 May 2023).

Vespro, C. (2006) ‘Stock Price and Volume Effects Associated with Compositional Changes in European Stock Indices’, *European Financial Management*, 12(1), pp. 103–127. doi: 10.1111/j.1354-7798.2006.00312.x.

Wang, C., Murgulov, Z. and Haman, J. (2015) ‘Impact of changes in the CSI 300 Index constituents’, *Emerging Markets Review*, 24, pp. 13–33. doi: 10.1016/j.ememar.2015.05.005.

Wurgler, J. and Zhuravskaya, E. (2002) ‘Does Arbitrage Flatten Demand Curves for Stocks?’, *The Journal of Business*, 75(4), pp. 583–608. doi: 10.1086/341636.

Xu, Y. (2017) ‘Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models’, *Political Analysis*, 25(1), pp. 57–76. doi: 10.1017/pan.2016.2.

Xu, Y. and Liu, L. (2022) ‘gsynth: Generalized Synthetic Control Method’, *R package version 1.2.1*. Available at: https://yiqingxu.org/packages/gsynth/ (Accessed: 10 March 2023).

Bibliography extension A

Table 3: Data sources

Variable	Source
Stock prices, Trading volume	Yahoo! Finance: (https://finance.yahoo.com) Verified by data from TradingView: (https://de.tradingview.com)
Market capitalization, P/B Ratio, EBITDA Margin, Dividend yield, Basic EPS	TradingView: (https://de.tradingview.com) Verified by data from Yahoo! Finance: (https://finance.yahoo.com)
Sector affiliation	Frankfurt stock exchange: (https://www.boerse-frankfurt.de)

Note: This table shows the main sources of the data used for the variables mentioned.
Source: Own table.

Bibliography extension B

Table 4: Data completion

Firm	Variable	Source
AFX	Basic EPS, EBITDA margin	Financial reports: https://www.zeiss.com/meditec-ag/investor-relations/reports-publications.html
AOX	Sector affiliation	Frankfurter Börse: https://www.boerse-frankfurt.de/aktie/alstria-office-reit-ag/unternehmensangaben
EVT	EBITDA margin	Macrotrends: https://www.macrotrends.net/stocks/charts/EVO/evotec-ag/ebitda-margin
G24	Basic EPS, EBITDA margin	Financial reports: https://www.scout24.com/en/investor-relations/financial-reports-presentations
HNR	EBITDA margin	Macrotrends: https://www.macrotrends.net/stocks/charts/HVRRY/hannover-ruck-se/ebitda-margin
O2D	Basic EPS, EBITDA margin	Financial reports: https://www.telefonica.de/investor-relations/publikationen/finanzpublikationen.html
WCH	Basic EPS, EBITDA margin	Financial reports: https://www.wacker.com/cms/en-us/about-wacker/investor-relations/financial-reports/financial-reports-overview.html

Note: This table shows the sources per company and variables that we use to complete the data. If multiple financial reports are used, the website to their download location is linked.
Source: Own table.

8 Appendix

Appendix A

Table 5: Literature overview

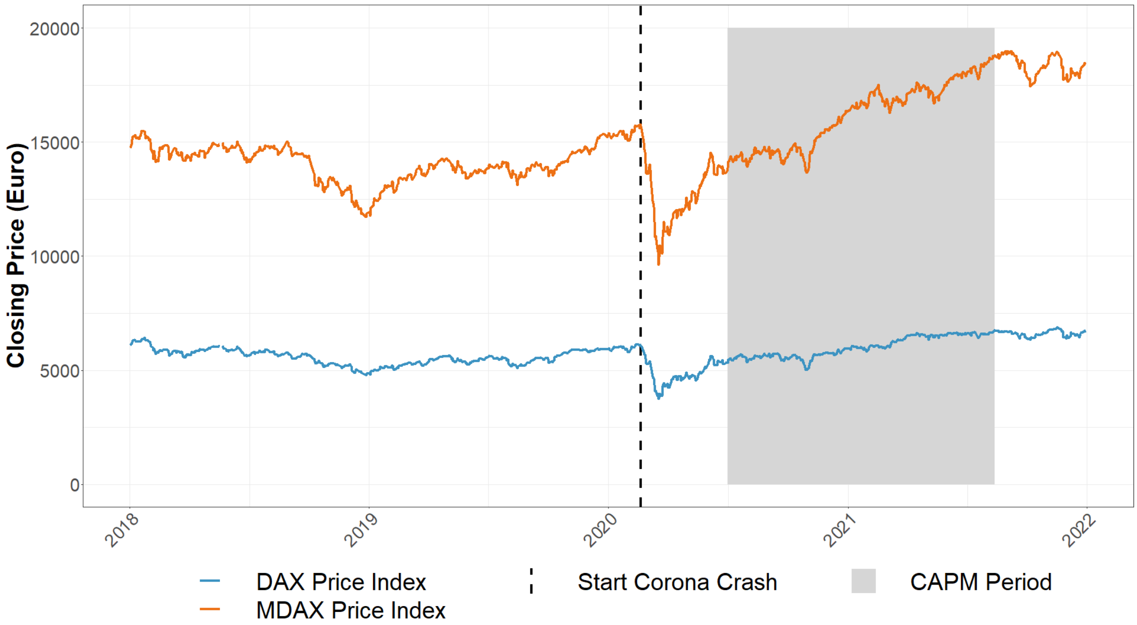
Paper	Hypothesis / Explanation	Index (Country)	Sample size Addition (Deletion)	Sample period
Harris and Gurel (1986)	Price pressure hypothesis	S&P 500 (USA)	194	1973-1983
Shleifer (1986)	Imperfect-substitutes hypothesis	S&P 500 (USA)	246	1966-1983
Jain (1987)	Information hypothesis	S&P 500 (USA)	110	1977-1983
Dhillon and Johnson (1991)	Information hypothesis, Imperfect-substitutes hypothesis	S&P 500 (USA)	187	1978-1988
Edmister, Graham and Pirie (1994)	Selection criteria hypothesis	S&P 500 (USA)	134	1983-1989
Beneish and Whaley (1996)	Imperfect-substitutes hypothesis, Price pressure hypothesis	S&P 500 (USA)	103	1986-1994
Lynch and Mendenhall (1997)	Imperfect-substitutes hypothesis, Price pressure hypothesis	S&P 500 (USA)	55 (53)	1976-1988
Wurgler and Zhuravskaya (2002)	Imperfect-substitutes hypothesis	S&P 500 (USA)	259	1976-1989
Denis et al. (2003)	Information effects, Imperfect substitutes hypothesis	S&P 500 (USA)	236	1987-1999
Elliott and Warr (2003)	Price pressure hypothesis, Imperfect-substitutes hypothesis	S&P 500 (USA)	187	1989-2000
Hegde and McDermott (2003)	Liquidity hypothesis, Imperfect-substitutes hypothesis	S&P 500 (USA)	74 (27)	1993-1998
Chen, Noronha and Singal (2004)	Investor awareness, Liquidity hypothesis	S&P 500 (USA)	760 (235)	1962-2000
Bechmann (2004)	Selection criteria hypothesis	KFX (Denmark)	52 (52)	1989-2001

Chakrabarti et al. (2005)	Imperfect substitutes hypothesis, Price pressure hypothesis, Liquidity hypothesis	29 different indices and countries	455	1998-2001
Becker-Blease and Paul (2006)	Liquidity hypothesis	S&P 500 (USA)	185	1980-2000
Elliot et al. (2006)	Investor awareness hypothesis, price pressure hypothesis	S&P 500 (USA)	147	1993-2001
Shankar and Miller (2006)	Price pressure hypothesis	S&P SmallCap 600 (USA)	556 (219)	1995-2002
Vespro (2006)	Price pressure hypothesis	CAC 40 & SBF 120 (France), FTSE 100 (UK)	24 (14) 23 (28)	1997-2001
Mazouz and Saadouni (2007)	Price pressure hypothesis	FTSE 100 (UK)	190 (187)	1984-2003
Cai (2007)	Information hypothesis	S&P 500 (USA)	427	1976-2001
Mase (2007)	Liquidity hypothesis, Price pressure hypothesis	FTSE 100 (UK)	132 (132)	1992-2005
Gygax and Otchere (2010)	Price pressure hypothesis Information hypothesis	S&P 500 (USA)	452 (187)	1987-2006
Chan, Kot and Tang (2013)	Information hypothesis, Liquidity hypothesis	S&P 500 (USA)	788 (244)	1962-2003
Wang, Murgulov and Haman (2015)	Price pressure hypothesis	CSI 300 (China)	266 (314)	2005-2012
Fernandes and Mergulhão (2016)	Imperfect substitutes hypothesis	FTSE 100 (UK)	138 (146)	1992-2010
Chen, Shiu and Wei (2019)	Investor awareness hypothesis	MSCI Standard indices from 38 Countries	1883 (1410)	2000-2015

Note: This table provides an overview of the papers relevant to us that explain the indexing effect with the help of the relevant hypotheses. It should be noted that the explanation "Information effects" means that the authors do not distinguish between the different information-based hypotheses. Further, the time period is often divided into different time periods and several studies are conducted. In each case, the "Hypothesis / Explanation" represents the most important findings of the paper.
Source: Own table. Data: See bibliography

Appendix B

Figure 7: Selection of the CAPM-period



Note: This chart shows the price of the DAX price index (blue) and the MDAX price index (orange) from 2018 to the end of 2021. The grey area represents our CAPM period. A strong price drop can be seen in both indices at the beginning of 2020, which was triggered by the COVID-19 pandemic. The dashed line marks the beginning of the corona crash. Such large price changes mark a major market anomaly.
Source: Own figure. Data: https://finance.yahoo.com

Appendix C

Table 6: Overview of treatment and control group

Number	Index Membership	Trading Symbol	Instrument	Group
1	MDAX	AIR	AIRBUS SE	Treatment
2	MDAX	BNR	BRENNTAG SE NA O.N.	Treatment
3	MDAX	HFG	HELLOFRESH SE INH O.N.	Treatment
4	MDAX	PUM	PUMA SE	Treatment
5	MDAX	QIA	QIAGEN NV EO -,01	Treatment
6	MDAX	SRT3	SARTORIUS AG VZO O.N.	Treatment
7	MDAX	SHL	SIEMENS HEALTH.AG NA O.N.	Treatment
8	MDAX	SY1	SYMRISE AG INH. O.N.	Treatment
9	MDAX	ZAL	ZALANDO SE	Treatment
10	MDAX	AIXA	AIXTRON SE NA O.N.	Control
11	MDAX	AOX	ALSTRIA OFFICE REIT-AG	Control
12	MDAX	AT1	AROUNDTOWN EO-,01	Control
13	MDAX	NDA	AURUBIS AG	Control
14	MDAX	BC8	BECHTLE AG O.N.	Control
15	MDAX	COK	CANCOM SE O.N.	Control
16	MDAX	AFX	CARL ZEISS MEDITEC AG	Control
17	MDAX	CBK	COMMERZBANK AG	Control
18	MDAX	COP	COMPUGROUP MED. NA O.N.	Control
19	MDAX	EVD	CTS EVENTIM KGAA	Control
20	MDAX	DUE	DUERR AG O.N.	Control
21	MDAX	EVK	EVONIK INDUSTRIES NA O.N.	Control
22	MDAX	EVT	EVOTEC SE INH O.N.	Control
23	MDAX	FRA	FRAPORT AG FFM.AIRPORT	Control

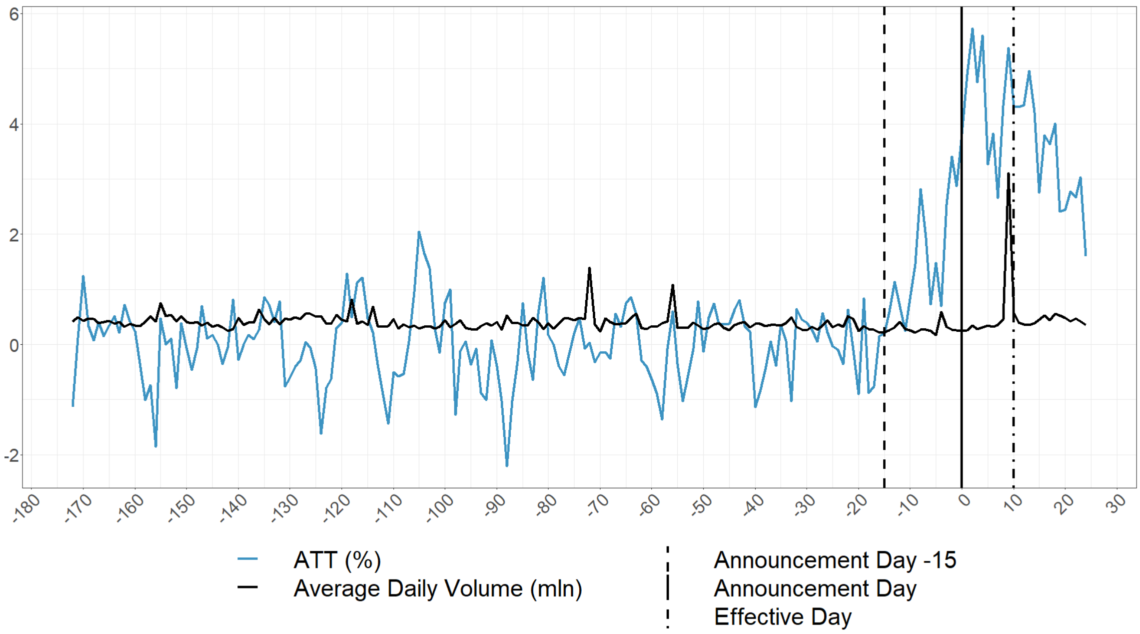
24	MDAX	FNTN	FREENET AG NA O.N.	Control
25	MDAX	FPE3	FUCHS PETROLUB VZO NA ON	Control
26	MDAX	G1A	GEA GROUP AG	Control
27	MDAX	GXI	GERRESHEIMER AG	Control
28	MDAX	GYC	GRAND CITY PROPERT.EO-,10	Control
29	MDAX	HNR1	HANNOVER RUECK SE NA O.N.	Control
30	MDAX	HLE	HELLA GMBH+CO. KGAA O.N.	Control
31	MDAX	BOSS	HUGO BOSS AG NA O.N.	Control
32	MDAX	SDF	K+S AG NA O.N.	Control
33	MDAX	KGX	KION GROUP AG	Control
34	MDAX	KBX	KNORR-BREMSE AG INH O.N.	Control
35	MDAX	LXS	LANXESS AG	Control
36	MDAX	LEG	LEG IMMOBILIEN SE NA O.N.	Control
37	MDAX	NEM	NEMETSCHEK SE O.N.	Control
38	MDAX	PSM	PROSIEBENSAT.1 NA O.N.	Control
39	MDAX	RAA	RATIONAL AG	Control
40	MDAX	RHM	RHEINMETALL AG	Control
41	MDAX	G24	SCOUT24 AG NA O.N.	Control
42	MDAX	SOW	SOFTWARE AG NA O.N.	Control
43	MDAX	SAX	STROEER SE + CO. KGAA	Control
44	MDAX	TEG	TAG IMMOBILIEN AG	Control
45	MDAX	TMV	TEAMVIEWER AG INH O.N.	Control
46	MDAX	O2D	TELEFONICA DTLD HLDG NA	Control
47	MDAX	TKA	THYSSENKRUPP AG O.N.	Control
48	MDAX	UN01	UNIPER SE NA O.N.	Control
49	MDAX	UTDI	UTD.INTERNET AG NA	Control
50	MDAX	VAR1	VARTA AG O.N.	Control

51	MDAX	WCH	WACKER CHEMIE O.N.	Control
52	MDAX	LHA	LUFTHANSA AG VNA O.N.	Excluded from the control group: Lack of information
53	MDAX	AG1	AUTO1 GROUP SE INH O.N.	Excluded from the control group: Enters MDAX in the training period
54	MDAX	BEI	BEIERSDORF AG O.N.	Excluded from the control group: Enters MDAX in the training period
55	MDAX	ECV	ENCAVIS AG INH. O.N.	Excluded from the control group: Drops to the SDAX in the treatment period
56	MDAX	HOT	HOCHTIEF AG	Excluded from the control group: Drops to the SDAX in the treatment period
57	MDAX	MOR	MORPHOSYS AG O.N.	Excluded from the control group: Drops to the SDAX in the treatment period
58	MDAX	NDX1	NORDEX SE O.N.	Excluded from the control group: Drops to the SDAX in the treatment period
59	MDAX	SAE	SHOP APOTHEKE EUROPE INH.	Excluded from the control group: Drops to the SDAX in the treatment period
60	MDAX	PAH3	PORSCHE AUTOM.HLDG VZO	Excluded from the treatment group: Enters MDAX in the training period

Note: This table is based on the index composition report of STOXX Ltd. At this time, all our analyzed companies are still in the MDAX. Furthermore, the table shows the trading symbol and the full name of the units. Based on further index reports, we assign the units to treatment and control group and describe why some are excluded from the control and treatment group.
Source: Own table. Data: Index Composition Report – MDAX (September 3,2021), https://www.dax-indices.com/de/web/dax-indices/zusammensetzung

Appendix D

Figure 8: ATT (%) compared to trading volume (mln) in the long run



Note: This figure compares the average number of shares sold or bought on a trading day in million for the treatment group (black) with the calculated ATT (blue) in percent over the whole period of 197 days. The trading volume is the equally weighted average of the volume, while the ATT was calculated using the GSC method. The y-axis applies to the ATT and the daily trading volume.
Source: Own figure. Data: https://finance.yahoo.com

Declaration of Authorship

for the master thesis within the master study of economics and management.

I hereby declare that the thesis with the title

“Analyzing the Indexing Effect on Stock Returns – Evidence from the German Stock Market”

has been composed by myself autonomously and no means other than those declared were used. I also declare that this thesis has not been handed in or published before in the same or similar form.

Furrer Jonas, 17-452-053

Cham, 17.05.2023 Signature:____________________________